

```

Db 70 NFFLLSLALADMFLGLLVPLSTIRSVESCWFFGDFCLRLHTYLDTLFCLTSIFHLFCIS 129
QY 186 LNDKSEVLSVLYQMATTTEVLSLQKLLAFAGLSVLVLTGLFMRFGLPCGWKXENI 245
Db 130 IDRHCAICDPLLYPSKFTVR-----ALRVLAGWG-----PAA--YTSL 168
QY 246 YITRQFQVDERERHQRPCV---LPLNKEERKXYIIPFWPTPKERNLGLFELPIL 301
Db 169 FLYTDVVEFLSOWLEMPGVCQQLLNK-----FWG-----LNFPLFFVPC 213
QY 302 IHLICWVLEAAVDYLLYRLIFSYSKQPSLPGFEVHLKHEKQGTQDIIHDSFNISVF 361
Db 214 IMISLVKIFVATRAQAQITLTK---SLAGAAKHERKAATLG----- 255
QY 362 EPNICPKPPELLSETWVPLSVIILVWLGLLSILMQLKILVSASYPSPVERKRIQYLH 421
Db 256 ----IVVGIYLL--CWLPTFTIDMVDLSLHFTTPPLV-FDIFWFAVENSACNPIIVFS 308
QY 422 AKLLKRSKOPGEG---VKRRLSLY 443
Db 309 YQWFRKALKTLTSQKVFSPOTRTVDLY 335

```

RESULT 2

```

US-09-370-253-6
; Sequence 6, Application US/09370253
; Patent No. 6165792
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Sakai, Hajime
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Amino Acid Transporters
; FILE REFERENCE: BB-1200
; CURRENT APPLICATION NUMBER: US/09/370,253
; EARLIER FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: 60/097,222
; EARLIER FILING DATE: August 20, 1998
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-370-253-6

```

```

Query Match 3.8%; Score 92; DB 4; Length 447;
Best Local Similarity 21.6%; Pred. No. 0.16;
Matches 55; Conservative 34; Mismatches 88; Indels 78; Gaps 12;

QY 15 WEIYSPRSGWMDFTQHLGVCGLVALISVGLLS-----VAACWPLSPSIIAAAASWIITCV 70
Db 26 WLPITSSRNAKWYSAPF---NVTAMVGAGVLSLPYAMSELGWPGLVILLSWITLY 81
QY 71 LLC-----CSKHACRILLVFLSCGLREG-----RNALIAAGTGI 105
Db 82 TLQWVEMHEMVPKRFDRYHELQHA-----FGEKGLGIWVVPQOLVVEGVNI 131
QY 106 VIL---GHVENIFHNFKGLDGWTC-NLRKSF-----SIHFPLK-KYIEAIOIYGLA 155
Db 132 VYMTGKSLKFKHDV--LCEGHGCKNIKUTYIMFIMFVHFLVLSQLPNFSISGVSLAA 189
QY 156 TPLSVFDDLVSNWQNTLAVLSFSPSHVLEAQLNDSKGEVLSVLYOMATTEV-----LSS 209
Db 190 AVMSLSYSTIANGASV-----DKGRVADVHLRATTSTGKVFGEFSA 232
QY 210 LGOKLLAFAGLSVL 224
Db 233 LGDVAFAAYAGHNVVL 247

```

RESULT 3

US-08-404-531B-29

```

; Sequence 29, Application US/08404531B
; Patent No. 5863724
; GENERAL INFORMATION:
; APPLICANT: Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson, Pamela
; APPLICANT: Thomas, Gilbert Cote, and Robert Gagel
; TITLE OF INVENTION: Sequence Encoding Mammalian Sulfonyleurea Receptor
; Patent No. 5863724
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESS: No. 5863724ris
; STREET: One Liberty Place 46th. Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/404,531B
; FILING DATE: 15-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; REFERENCE/DOCKET NUMBER: BYLR-0003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1498 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-404-531B-29

```

```

Query Match 3.7%; Score 91; DB 2; Length 1498;
Best Local Similarity 18.6%; Pred. No. 1.3;
Matches 96; Conservative 75; Mismatches 168; Indels 176; Gaps 25;

```

```

QY 65 WITCVLLCCSKHARCFILLV-----FLSCGLREGRNALIAAGTGIVILGHVENI--EHN 117
Db 75 WILTFILL-----FVLVCEIAGILSDGVTESRHLHLYMPAGMAFMAAITSVYVYHN 126
QY 118 FKGLLDGMTCLNLRKASFSIHFPPLKKYIEAIOIYGLATP---LSVFDDLVSNWQNT--- 170
Db 127 IE-----TSNFPKLLIAL-LIYWTLAFTIKTKFYKFDHAGFSQLRFC 170
QY 171 ---LAVLSFSPSHVLEAQLNDSKGEVLSVLYOMATTEVLSLQKLLAFAGLSVLVLTGT 227
Db 171 LTGLVILVYGLMLLVEVNVIRVRYIFFTKPREVKPPEDLDLGVRFLO--PFVNLLSKGT 229
QY 228 GLEMKRFLG-----PCGMKYENIYITRQFQVDERERHQ--QRP----- 264
Db 230 YWMNNAFKTAHKPIDLRATAKLPAMRALNY-ORLCVAFDAQARKDTQSPQAGARA 288
QY 265 -----CVL-----PLNKEERKXYIIPFWPTPKER 290
Db 289 RALCHAFGRLLSSTFRILADLLGAGPLCIFGIVDHLGKENH-----VFQP---KT 338
QY 291 KNLGLFF-----LPILHLCIWLVAFAVDYLLYRLIFSYSKQPSLPGFEVHL 338
Db 339 QFLGVYFVSSQEFGLNAYVAVLLEFLAL-----LLQRTFLQASYVV---ALETGI 385
QY 339 KLHGEKQGT--QDIHDSFNISV-----FEPNCIPKPKFLLSETW--VPLSV 382
Db 386 NLRGAIQTKIYNKIMHMTSNLSMGEMTAGTCNLVAIDTNQLMWWFFFLCPNLWTMPVQI 445

```

Result No.	Query			ID	Description	
	Score	Match	Length			
1	93	3.8	337	2	US-08-467-559B-2	Sequence 2, Appli
2	92	3.8	447	4	US-09-370-253-6	Sequence 6, Appli
3	91	3.7	1498	2	US-08-400-531B-29	Sequence 29, Appli
4	91	3.7	1498	3	US-08-476-900A-29	Sequence 29, Appli
5	91	3.7	1498	3	US-08-488-546A-29	Sequence 29, Appli
6	91	3.7	1582	2	US-08-404-531B-9	Sequence 9, Appli
7	91	3.7	1582	3	US-08-476-900A-9	Sequence 9, Appli
8	91	3.7	1582	3	US-08-488-546A-9	Sequence 9, Appli
9	91	3.7	1582	4	US-08-726-320-5	Sequence 5, Appli
10	91	3.7	1582	4	US-09-208-716-5	Sequence 5, Appli
11	89	3.6	554	4	US-08-180-371-2	Sequence 2, Appli
12	88.5	3.6	2237	1	US-08-455-543A-48	Sequence 48, Appli
13	88.5	3.6	2237	2	US-08-223-305C-48	Sequence 48, Appli
14	88.5	3.6	2296	2	US-08-286-819A-27	Sequence 27, Appli
15	88.5	3.6	2296	3	US-08-980-357-27	Sequence 27, Appli
16	88.5	3.6	2337	3	US-08-713-118-2	Sequence 2, Appli
17	88.5	3.6	2337	4	US-09-452-007-2	Sequence 2, Appli
18	88.5	3.6	2339	1	US-08-455-543A-47	Sequence 47, Appli
19	88.5	3.6	2339	2	US-08-223-305C-47	Sequence 47, Appli
20	87.5	3.6	447	4	US-09-370-253-10	Sequence 10, Appli
21	87.5	3.6	823	1	US-07-745-206A-15	Sequence 15, Appli
22	87.5	3.6	823	2	US-08-311-363-15	Sequence 15, Appli
23	87.5	3.6	1754	1	US-07-745-206A-13	Sequence 13, Appli
24	87.5	3.6	1754	2	US-08-311-363-13	Sequence 13, Appli
25	87.5	3.6	1835	3	US-08-836-325-15	Sequence 15, Appli
26	87	3.6	394	2	US-08-742-440A-8	Sequence 8, Appli
27	87	3.6	453	1	US-08-570-157-7	Sequence 7, Appli


```
Db 75 WILTFILL-----FVLCEIAGILSDGVTESRHLHLYMPAGMAPMAAITSVVYHN 126
Qy 118 FKGLDGMTNLRKSFISHPPLKKYIEATQIWIYGLATP-----LSVFDLVSNNQOT--- 170
Db 127 IE-----TSNEPKLLIAL-LIYWTAFITTKIKFVKFYDHAIGFSQRLRFC 170
Qy 171 ---LAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEVLSLQKLLAFAGLSLVLLGT 227
Db 171 LTGLLVILYGMILLVEVNVIRVRYIFKTPREVKKPPEDLDLGVRFLO-PFVNLLSKGT 229
Qy 228 GLFMKRFGL-----PCGWKYENIYITROFQVDERERHQ-ORP----- 264
Db 230 YWMNAFIKTAHKPIDLRAIAKPIAMRALTNY-ORLCVAFDAQARKDTOSPOCARAIW 288
Qy 265 -----CVL-----PLNKEERRKYIITFTWTPKER 290
Db 289 RALCHAFGRRLILSSTFRILADLLGFAGPLCIFGIVDHLGKENH-----VFQP---KT 338
Qy 291 KNGLGFF-----LPILHLCIWLFAAVDYLLYRLIFSVMKQFOSLPGEVHL 338
Db 339 QFLGVYFVSSQEFNGAYVLAVALFLAL-----LQRTFLQASYV---AIETGI 385
Qy 339 KLHGEKQGT--QDIHDSSENIW-----FEPNCIPKPKFLLSETW-VPLSV 382
Db 386 NLRGAIQTKIYNKIMHSTSNLSMGENTAGQICNLVAIDTNLMWFFFLCPNLWTPVOI 445
Qy 383 I---LLILVMLGLSSILMOLKILVSASFYPSVERKRIQYLHAKLLKRSKQPLGEVKRR 439
Db 446 IGVVILLYILGVSALIGAIVILLAP-----VQYFVATKLSQAQRTTLEHSNER 495
Qy 440 L-----SLYTKIHFWLPV-----LKMIRKKOM 462
Db 496 LKQTNEMLRGMKLLKYAWESIFCSRVEVTRKEM 530
```

RESULT 6

```
US-08-404-531B-9
; Sequence 9, Application US/08404531B
; Patent No. 5863724
; GENERAL INFORMATION:
; APPLICANT: Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson, Pamela
; APPLICANT: Thomas, Gilbert Cote, and Robert Gagel
; TITLE OF INVENTION: Sequence Encoding Mammalian Sulfonylurea Receptor
; Patent No. 5863724
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: No. 5863724ris
; STREET: One Liberty Place 46th. Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/404,531B
; FILING DATE: 15-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; REFERENCE/DOCKET NUMBER: BYLR-0003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1582 amino acids
```

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-404-531B-9
```

Query Match 3.7%; Score 91; DB 2; Length 1582;

```
Best Local Similarity 18.8%; Pred. No. 1.4;
Matches 96; Conservative 75; Mismatches 168; Indels 176; Gaps 25;
```

```
Qy 65 WILTCVLLCCSKHARCFILLV-----FLSCGLREGRNALIAAGTGIVILGHVENI--PHN 117
Db 75 WILTFILL-----FVLCEIAGILSDGVTESRHLHLYMPAGMAPMAAITSVVYHN 126
Qy 118 FKGLDGMTNLRKSFISHPPLKKYIEATQIWIYGLATP-----LSVFDLVSNNQOT--- 170
Db 127 IE-----TSNEPKLLIAL-LIYWTAFITTKIKFVKFYDHAIGFSQRLRFC 170
Qy 171 ---LAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEVLSLQKLLAFAGLSLVLLGT 227
Db 171 LTGLLVILYGMILLVEVNVIRVRYIFKTPREVKKPPEDLDLGVRFLO-PFVNLLSKGT 229
Qy 228 GLFMKRFGL-----PCGWKYENIYITROFQVDERERHQ-ORP----- 264
Db 230 YWMNAFIKTAHKPIDLRAIAKPIAMRALTNY-ORLCVAFDAQARKDTOSPOCARAIW 288
Qy 265 -----CVL-----PLNKEERRKYIITFTWTPKER 290
Db 289 RALCHAFGRRLILSSTFRILADLLGFAGPLCIFGIVDHLGKENH-----VFQP---KT 338
Qy 291 KNGLGFF-----LPILHLCIWLFAAVDYLLYRLIFSVMKQFOSLPGEVHL 338
Db 339 QFLGVYFVSSQEFNGAYVLAVALFLAL-----LQRTFLQASYV---AIETGI 385
Qy 339 KLHGEKQGT--QDIHDSSENIW-----FEPNCIPKPKFLLSETW-VPLSV 382
Db 386 NLRGAIQTKIYNKIMHSTSNLSMGENTAGQICNLVAIDTNLMWFFFLCPNLWTPVOI 445
Qy 383 I---LLILVMLGLSSILMOLKILVSASFYPSVERKRIQYLHAKLLKRSKQPLGEVKRR 439
Db 446 IGVVILLYILGVSALIGAIVILLAP-----VQYFVATKLSQAQRTTLEHSNER 495
Qy 440 L-----SLYTKIHFWLPV-----LKMIRKKOM 462
Db 496 LKQTNEMLRGMKLLKYAWESIFCSRVEVTRKEM 530
```

RESULT 7

```
US-08-476-900A-9
; Sequence 9, Application US/08476900A
; Patent No. 6031150
; GENERAL INFORMATION:
; APPLICANT: Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson
; TITLE OF INVENTION: Sequence Encoding Mammalian Sulfonylurea Receptor
; Patent No. 6031150
; TITLE OF INVENTION: and Method of Detecting Persistent Hyperinsulinemic Hypoglycemia
; TITLE OF INVENTION: Infancy
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: No. 6031150ris
; STREET: One Liberty Place 46th. Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,900A
```


FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
REFERENCE/DOCKET NUMBER: BYLR-0027
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1582 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-900A-9

Query Match 3.7%; Score 91; DB 3; Length 1582;
Best Local Similarity 18.6%; Pred. No. 1.4;
Matches 96; Conservative 75; Mismatches 168; Indels 176; Gaps 25;

QY 65 WIITCVLLCCSKHARCFILLV-----FLSCGLREGRNALIAAGTGIVILGHVENI--FHN 117
DB 75 WILTFILL-----FVLVCEIAEGILSDGVTSERHLLHLYMPAGMAFMAAITSVYVYHN 126
QY 118 FKGLDGMTCNLRKSFHSIHFPFLKKYIEAOWIYGLATP-----LSVFDLVSWNQT--- 170
DB 127 IE-----TSNFPKLLIAL-LIYWTAFITKTKFVKFYDHAIGFSQLRFC 170
QY 171 ---LAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEVLSLQKLLAFAGLSVLVLTG 227
DB 171 LTGLLVILYGMILLVEVNVIRVRRYIFFKTPREVKPPEDLDLQGVRFLO-PFVNLLSKGT 229
QY 228 GLFMKRFGL-----PCGWKENIYITRQFVQFDERHQ-QRP----- 264
DB 230 YWMNAFIKTAHKPIDLRAIAKLPAMRALTNY-QLCVAFDAQARKDTQSPQGARAIW 288
QY 265 -----CVL-----PLNKEERRYVIITPTWTPKER 290
DB 289 RALCHAFGRRLILSSFRILADLLGAGPLCIFGIVDHLGKENH-----VFQP---KT 338
QY 291 KNLGLFF-----LPILHLCIWLFAAVDYLLYRLIFSQKQPLGFEVHL 338
DB 339 QFLGVYFVSSQEFGLNAYVLAIVLLAL-----LQRTFLOASYV---ALETGI 385
QY 339 KLHGEKQGT-QDIHDSFNSV-----FPCWIKPKFLLSETW-VPLSV 382
DB 386 NURGAITQKIYKIMHSTNSLMGEMTAGQICNLVAIDTNQLMWFFFLCPNLWTPVQI 445
QY 383 I---LLILVMLGLSSILMQKILVSASFVPSVERKRIQYLHAKLLKRSKQPLGEVKRR 439
DB 446 IVGVILLYILGVSAIGAAVILLAP-----VQYFVATKLSQAQRTTLEHSNER 495
QY 440 L-----SLYLTKIHFVLPV-----LKMIRKKOM 462
DB 496 LKOTNEMLRGMKLLKYAWESIFCSRVEVTRRKM 530

RESULT 8
US-08-488-546A-9
Sequence 9, Application US/08488546A
Patent No. 6054313
GENERAL INFORMATION:
APPLICANT: Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson, Pamela
APPLICANT: Thomas, Gilbert Cole, and Robert Gagel
TITLE OF INVENTION: Sequence Encoding Mammalian Sulfonyleura Receptor
Patent No. 6054313
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: No. 6054313ris
STREET: One Liberty Place 46th. Floor

CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,546A
FILING DATE: 07-JUNE-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/404,531
FILING DATE: 15-MARCH-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
REFERENCE/DOCKET NUMBER: BYLR-0026
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1582 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-488-546A-9

Query Match 3.7%; Score 91; DB 3; Length 1582;
Best Local Similarity 18.6%; Pred. No. 1.4;
Matches 96; Conservative 75; Mismatches 168; Indels 176; Gaps 25;
QY 65 WIITCVLLCCSKHARCFILLV-----FLSCGLREGRNALIAAGTGIVILGHVENI--FHN 117
DB 75 WILTFILL-----FVLVCEIAEGILSDGVTSERHLLHLYMPAGMAFMAAITSVYVYHN 126
QY 118 FKGLDGMTCNLRKSFHSIHFPFLKKYIEAOWIYGLATP-----LSVFDLVSWNQT--- 170
DB 127 IE-----TSNFPKLLIAL-LIYWTAFITKTKFVKFYDHAIGFSQLRFC 170
QY 171 ---LAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEVLSLQKLLAFAGLSVLVLTG 227
DB 171 LTGLLVILYGMILLVEVNVIRVRRYIFFKTPREVKPPEDLDLQGVRFLO-PFVNLLSKGT 229
QY 228 GLFMKRFGL-----PCGWKENIYITRQFVQFDERHQ-QRP----- 264
DB 230 YWMNAFIKTAHKPIDLRAIAKLPAMRALTNY-QLCVAFDAQARKDTQSPQGARAIW 288
QY 265 -----CVL-----PLNKEERRYVIITPTWTPKER 290
DB 289 RALCHAFGRRLILSSFRILADLLGAGPLCIFGIVDHLGKENH-----VFQP---KT 338
QY 291 KNLGLFF-----LPILHLCIWLFAAVDYLLYRLIFSQKQPLGFEVHL 338
DB 339 QFLGVYFVSSQEFGLNAYVLAIVLLAL-----LQRTFLOASYV---ALETGI 385
QY 339 KLHGEKQGT-QDIHDSFNSV-----FPCWIKPKFLLSETW-VPLSV 382
DB 386 NURGAITQKIYKIMHSTNSLMGEMTAGQICNLVAIDTNQLMWFFFLCPNLWTPVQI 445
QY 383 I---LLILVMLGLSSILMQKILVSASFVPSVERKRIQYLHAKLLKRSKQPLGEVKRR 439
DB 446 IVGVILLYILGVSAIGAAVILLAP-----VQYFVATKLSQAQRTTLEHSNER 495
QY 440 L-----SLYLTKIHFVLPV-----LKMIRKKOM 462
DB 496 LKOTNEMLRGMKLLKYAWESIFCSRVEVTRRKM 530

QY 118 FKGLDGMTCLNRAKSFSTHPLKKYIEAIOIYGLATP-----LSVFDDLYSWNOT--- 170
Db 127 IE-----TSNPKLLIAL-LIWTIAFITKTKFKYFHAIGFSLREC 170
QY 171 ---LAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTEVLSLGGKLLAFAGLSVLGGT 227
Db 171 LTGLLVILYGMLLLEVNVIRVRRYIEFKTPREVKKPPDLDLGVRFLO-PFVNLLSKGT 229
QY 228 GLFMRFLG-----PCGWKENIYITRFVQFDERHO-QRP----- 264
Db 230 YWMNAFIKTAHKPIDLRAIAKLPIAMRALTNY-ORLCVAFDAQARKDTQPOGARAIW 288
QY 265 -----CVL-----PLNKEERRKYIIPFTWPTPKR 290
Db 289 RALCHAFORRLILSTFRILADLLGAFAGPLCIFGVHDLGKHNH-----VFQP-----KT 338
QY 291 KNLGLFF-----LPLIHLCIWLFAAVDYLLYRIFSVSKQFSLPGFEVHL 338
Db 339 QFLGVYFVSSQFELGNAYVLAVALLEL-----LLQRTFLQASYV-----ALETGI 385
QY 339 KLHGKQGT--ODIHDSSFNISV-----PEPNCIPKPKFLLSETW-VPLSV 382
Db 386 NLRGAIQTKIYNKIMHSTSNLSMGEMTAGQICNLVAIDTNOLMWFELCPNLWTMPVOI 445
QY 383 I---LLILVMLGLSSILMOLKILVSAFYPVSVERKRIQYLHAKLLKSKQPLGEVKR 439
Db 446 IGVVILYILGVSALICAAVILLAP-----VOYFVATKLSQAORTTLEHSNER 495
QY 440 L-----SLYLTKIHFPLPV-----LKMIRKKOM 462
Db 496 LKQTNEMLRGMKLLLYAWESIFCSRVEVTRKEM 530

RESULT 11

US-08-180-371-2

; Sequence 2, Application US/08180371

; Patent No. 6254861

; GENERAL INFORMATION:

; APPLICANT: Choudhury, Chandra

; TITLE OF INVENTION: Hematopoietic Growth Factor Derived

; FROM T Lymphocytes and Methods of Use Therefor

; NUMBER OF SEQUENCES: 18

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &

; STREET: 6300 Sears Tower, 233 South Wacker Drive

; CITY: Chicago

; STATE: Illinois

; COUNTRY: United States of America

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/180,371

; FILING DATE: 12-JAN-1994

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 906866

; FILING DATE: 01 July 1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Gruber, Lewis S.

; REGISTRATION NUMBER: 30,060

; REFERENCE/DOCKET NUMBER: 27620/31668

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (312) 474-6300

; TELEFAX: (312) 474-0448

; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 554 amino acids

;

;

;

;

;

;

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-180-371-2

Query Match 3.6%; Score 89; DB 4; Length 554;

Best Local Similarity 20.7%; Pred. No. 0.47; Indels 132; Gaps 19;

Matches 80; Conservative 46; Mismatches 129;

QY 133 SFSIHFPLKKYIE---AIQMIYGLATPLSYF-----DDLV----- 165

Db 36 SSISSIFYSGREYKNDSESTQCYADTLPLSYLHVLSNLVIKQPRQROIYSQCQFGPSQ 95

QY 166 SWNOTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTEVLSLGGKLLAFAG----- 219

Db 96 SSGKRLPSSFLTCSHICHQSLSDESMTYQAEKDEQDTLGLSSCSLDVLCIGGIRRESF 155

QY 220 -LSIVLL-----GTCLFMKRELPGCWKENIYITRFVQFDERERHOQRPVL 267

Db 156 PLIYVMLTHRSHSTARLKGITHLASRF-----SSSTVFYSKVF-----OACLL 198

QY 268 PLNKEERRKVIIPFTWPTPKERNLGLFFL--PLIHLCIWLFAAVDYLLYRIFSVS 325

Db 199 -LIKLFYILLIIPFTWFPQSSMSLSIICLSKPLKIITKTWIFQTLDF--QRL----- 250

QY 326 KQFQSLPGEFVHLKHEKQGTQDIHDSSFNISVFEPNCIPKPKFLLSETWVPLSVILL 385

Db 251 -----LPTW--HLQLN-----IQOTAPS-----QSNSTQRLIFLSAVFVRKGSLLA 290

QY 386 I-----LVNMLGLSSILMOLKILVSAFSPSVF-----RKRIQY---LHAKLL-- 425

Db 291 TSPFKSLRLQTLQSDLM-----GTFFFPKTSVTSHLDSCLNFKKHHSYAYSFHAQLIFP 344

QY 426 -----KRSKOPLGVEVKR 438

Db 345 XQNTQGIILGRVMNVNSAKPTVTEKK 371

RESULT 12

US-08-455-543A-48

; Sequence 48, Application US/08455543A

; Patent No. 5792846

; GENERAL INFORMATION:

; APPLICANT: Harpold, Michael

; APPLICANT: Ellis, Steven

; APPLICANT: Williams, Mark

; APPLICANT: Feldman, Daniel

; APPLICANT: McCue, Ann

; APPLICANT: Brenner, Robert

; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND

; NUMBER OF SEQUENCES: 57

; CORRESPONDENCE ADDRESS:

; ADDRESS: Brown, Martin, Haller & McClain

; STREET: 1660 Union Street

; CITY: San Diego

; STATE: California

; COUNTRY: USA

; ZIP: 92101-2926

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/455,543A

; FILING DATE: May 31, 1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/223,305

; FILING DATE: April 4, 1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/868,354

;

;

;

;

;

Query Match 3.6%; Score 88.5; DB 1; Length 2237;
Best Local Similarity 19.0%; Pred. No. 4.5;
Matches 84; Conservative 64; Mismatches 126; Indels 169

```
QY   63 ASWIIITCVLCCSKH-----ARCFILLVFLSCGLREGRNALIAAGTIVI 107  
    ! : ! : ! : ! : ! : ! : ! :  
Db   103 ATTIANGIVLALEOHLDPDGKTPMSERLDDPEFYIGF-----CFEAGIKIUA 151  
    ! : ! : ! : ! : ! : ! : ! :  
QY   108 LGHYENIFHNFKGILGDGMTCNLRAKSFSIHPPLKKYIEAQIWYGATPLSVDDLVSW 167  
    ! ! ! ! ! : ! : ! : ! : ! : ! :  
Db   152 LGFV--FHKGSYLRGNWN-----VNDFVVLTGI--LATAGTDFFD----- 187  
    ! : ! : ! : ! : ! : ! : ! :  
QY   168 NQTL-AVSLSPSHSVAEQLNDSKEVLSSVLYOMATTTEVLSSLGOKLLAFAGLSVLLG 226  
    ! ! ! : ! : ! : ! : ! : ! : ! :  
Db   188 LRTLRAVRVLRPLKLVSQ-IPSLQOVWLKSIMKAMVPPLLQI-----GULLFPAILMFAGI 241  
    ! : ! : ! : ! : ! : ! : ! :  
QY   227 TGLFMKRF-----LG--PCGWKYENIYTQQVFDERERHQORPCVLPUNK 271  
    ! : ! : ! : ! : ! : ! : ! :  
Db   242 LEFTMGFAHFHKACPNSTDAEPVGDFPGC-----KEAPARLCBGD 280  
    ! : ! : ! : ! : ! : ! : ! :  
QY   272 EERRKYVIPTFWPTKERNMLGLFFLPILTHLCIWLFAAVDYLLRYLFPSVKQCOSL 331  
    ! ! ! ! ! : ! : ! : ! : ! : ! :  
Db   281 TECREX-----WGPQ---NFGITNE-----NILFAILTVFOCI 311  
    ! : ! : ! : ! : ! : ! : ! :  
QY   332 PGFEVHLKLHGEGKTODITHDSSFNISVPEPNCIPPKFLLSTWTWPLSUILILV--- 388  
    ! : ! : ! : ! : ! : ! : ! :  
Db   312 -----TMEGWDILYNTN-----DAAGTNWNLVPIPLIITGSF 345  
    ! : ! : ! : ! : ! : ! : ! :  
QY   389 -MLGLSSIQLMWLKILYSASGYPSVERKRIOYLHAKLLKRKSGPLGEVBRRLSYLTKI 447  
    ! ! ! : ! : ! : ! : ! : ! : ! :  
Db   346 FMLNLVLGVL-----SGEF--AKERERVNRRAFKLRRQOQ-----IERELNGYLE-- 390  
    ! : ! : ! : ! : ! : ! : ! :  
QY   448 HFWLPVLKMTKRKKOMDMASADKS 470  
    ! : ! : ! : ! : ! : ! : ! :  
Db   391 --WIFFKAEEYNLMAEEDRNBEEKS 411
```

RESULT 13
 US-08-223-305C-48
 ; Sequence 48, Application US/08223305C
 ; Patent No. 5851824
 ; GENERAL INFORMATION:
 ; APPLICANT: Harpold, Michael
 ; APPLICANT: Ellis, Steven
 ; APPLICANT: Williams, Mark
 ; APPLICANT: Feldman, Daniel
 ; APPLICANT: McCue, Ann
 ; APPLICANT: Brenner, Robert
 ; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
 ; TITLE OF INVENTION: METHODS
 ; NUMBER OF SEQUENCES: 57
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Brown, Martin, Haller & McClain
 ; STREET: 1660 Union Street
 ; CITY: San Diego
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 92101-2926
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq Version 1.5
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/223,305C
 ; FILING DATE: April 4, 1994
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 07/868,354
 ; FILING DATE: April 10, 1992
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/745,206
 ; FILING DATE: 15-AUG-1991
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/620,250
 ; FILING DATE: 30-NOV-1990
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/482,384
 ; FILING DATE: 20-FEB-1990
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/603,751
 ; FILING DATE: 04-APR-1989
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: WO PCT/US89/01408
 ; FILING DATE: 04-APR-1989
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/176,899
 ; FILING DATE: 04-APR-1988
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Seidman, Stephanie L.
 ; REGISTRATION NUMBER: 33,779
 ; REFERENCE/DOCKET NUMBER: 52516 (P519739)
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619)238-0099
 ; TELEFAX: (619)238-0062
 ; INFORMATION FOR SEQ ID NO: 48:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2237 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; FRAGMENT TYPE: internal
 US-08-223-305C-48

Query Match 3.6%; Score 88.5; DB 2; Length 2237;
Best Local Similarity 19.0%; Pred. No. 4.5;
Matches 84; Conservative 64; Mismatches 126; Indels 169

ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,357
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,819
FILING DATE: 05-AUG-1994
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, NO. 6013508man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 2296 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-980-357-27

Query Match 3.6%; Score 88.5; DB 3; Length 2296;
Best Local Similarity 22.4%; Pred. No. 4.7;
Matches 68; Conservative 41; Mismatches 109; Indels 85; Gaps 13;
QY 191 GEVLVLYQMATTTEVLSSLGQKLLAFAGLSVLVLTG-----LPMKRELGPCGW 240
Db 2007 GHYLRALRQ-----DSLIIHRIARGGISHRPLSTGSSASLSAWSLMKIHLLH----- 2054
QY 241 KYENIVITROFQVDER---ERHQORPCVLP LNKEERRKYVIPTFWPTPKERNLGLFF 297
Db 2055 -----WIOGEIIDCNLRGKTAQSOTRLCL---RGFKYFILPTI--LRRLAKMKLFF 2103
QY 298 LPILHLCIWLFAVDYLLYRIFSVSKQFSLPGPEVHLKLGKEQGTQDIHDS-- 355
Db 2104 LLLLLL-----IYLGVDYVNEALFSQEKVEFQN---YDQNPKEHLENGTSENTQKTI 2156
QY 356 -----FNISVFEPC-----IPKPKFLLSETW----- 377
Db 2157 EEQVYQGNLLINSKYPVROEVQSISYLNMTNMTDGLIVIFICQKKHKNFORQWSMMLR 2216
QY 378 VPLSVILLIILVGLLSILLMQIKLVSAFSPYVERKRIQYLHAKLLKKRSKQPLGEVK 437
Db 2217 VALVILLIILVAIETLMSKVCFTKKWGLSMP-YQQVIVSIIOVYHMDQARKWN-EPLKESG 2274
QY 438 RRL 440
Db 2275 KKM 2277

Search completed: January 31, 2002, 08:02:24
Job time: 105 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2003, 13:25:47 ; Search time 50 Seconds
(without alignments)
1017.144 Million cell updates/sec

Title: US-09-713-098-2

Perfect score: 2448

Sequence: 1 MGWTSGTDIFLSLWEIVYS.....LPVLKMRKKQMDASADKS 470

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/ECT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/ECTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1799.5	73.5	352	10 US-09-789-561-98	Sequence 98, Appl
2	1314	53.7	291	10 US-09-789-561-142	Sequence 142, App
3	1309	53.5	257	10 US-09-789-561-140	Sequence 140, App
4	1309	53.5	257	10 US-09-789-561-141	Sequence 141, App
5	171	7.0	485	9 US-09-798-889-55	Sequence 55, Appl
6	100.5	4.1	729	9 US-10-205-823-367	Sequence 367, App
7	96	3.9	405	9 US-09-895-913A-234	Sequence 234, App
8	95	3.9	263	9 US-09-924-340-40	Sequence 40, Appl
9	95	3.9	263	9 US-09-992-600A-40	Sequence 40, Appl
10	95	3.9	263	9 US-10-000-489-40	Sequence 40, Appl
11	95	3.9	263	9 US-10-000-986-40	Sequence 40, Appl
12	93	3.8	337	9 US-10-180-946-2	Sequence 2, Appli
13	93	3.8	933	10 US-09-815-242-11817	Sequence 11817, A
14	92	3.8	1451	9 US-09-998-027-1	Sequence 1, Appli
15	92	3.8	1471	9 US-09-998-027-4	Sequence 4, Appli
16	91.5	3.7	581	10 US-09-835-270-2	Sequence 2, Appli
17	90.5	3.7	424	10 US-09-817-774-25	Sequence 25, Appl
18	90.5	3.7	489	9 US-10-146-835-4	Sequence 4, Appli
19	90	3.7	309	9 US-09-510-332-85	Sequence 85, Appl

20	89.5	3.7	463	10 US-09-815-242-10144	Sequence 10144, A
21	89	3.6	2336	12 US-10-033-026-10	Sequence 10, Appl
22	88.5	3.6	2237	12 US-10-033-026-8	Sequence 8, Appli
23	88.5	3.6	2339	12 US-10-033-026-6	Sequence 6, Appli
24	88.5	3.6	2343	12 US-10-033-026-4	Sequence 4, Appli
25	88	3.6	436	10 US-09-815-242-13218	Sequence 13218, A
26	87.5	3.6	734	9 US-09-738-626-4227	Sequence 4227, Ap
27	87.5	3.6	772	10 US-09-935-799A-2	Sequence 2, Appli
28	87.5	3.6	772	10 US-09-935-799A-5	Sequence 5, Appli
29	87.5	3.6	807	10 US-09-895-686-6	Sequence 6, Appli
30	87.5	3.6	986	9 US-10-120-604-101	Sequence 101, App
31	87.5	3.6	986	10 US-09-747-835A-53	Sequence 53, Appl
32	87.5	3.6	1083	10 US-09-789-481C-2	Sequence 2, Appli
33	87.5	3.6	1325	10 US-09-747-835A-24	Sequence 24, Appl
34	87.5	3.6	1346	10 US-09-747-835A-20	Sequence 20, Appl
35	87.5	3.6	1346	10 US-09-747-835A-61	Sequence 61, Appl
36	87.5	3.6	1503	9 US-10-007-706-1	Sequence 1, Appli
37	87.5	3.6	1835	9 US-09-457-571-15	Sequence 15, Appl
38	87	3.6	557	10 US-09-815-242-12536	Sequence 12536, A
39	86.5	3.5	2310	10 US-09-995-542-10	Sequence 10, Appl
40	86	3.5	750	9 US-09-746-783-86	Sequence 86, Appl
41	86	3.5	2273	10 US-09-995-542-12	Sequence 12, Appl
42	84.5	3.5	209	9 US-10-012-542-150	Sequence 150, App
43	84.5	3.5	1116	9 US-09-835-976B-12	Sequence 12, Appl
44	84.5	3.5	1377	9 US-10-205-342-25	Sequence 25, Appl
45	84	3.4	344	10 US-09-880-261-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-789-561-98
; Sequence 98, Application US/09789561
; Patent No. US20020064818A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 52 Human secreted proteins
; FILE REFERENCE: PZ043P1
; CURRENT APPLICATION NUMBER: US/09/789,561
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: PCT/US00/24008
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,317
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/152,315
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 98
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (284)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-789-561-98

Query Match	73.5%;	Score 1799.5;	DB 10;	Length 352;
Best Local Similarity	96.4%;	Pred. No. 2e-160;		
Matches 344;	Conservative 2;	Mismatches 4;	Indels 7;	Gaps 1;
QY	27	MDFIQHLGVCCVLVALISVGLLSVAACWFLPSIIIAAASWIIITCVLLCCSKHARCFILLVF	86	
Db	1	MDFIQHLGVCCVLVALISVGLLSVAACWFLPSIIIAAASWIIITCVLLCCSKHARCFILLVF	60	
QY	87	LSGCLREGNALIAAGTGIVILGHVENIHFNFKGLDGMTCNLRKSFSTHFPPLKKYIE	146	
Db	61	LSGCLREGNALIAAGTGIVILGHVENIHFNFKGLDGMTCNLRKSFSTHFPPLKKYIE	120	
QY	147	AIQWTYGLATPLSVDFDVLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLVLYOMATTEV	206	

Db 121 AIQWIYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 180
QY 207 LSSLGOKLLAFAGLSVLVLTGTLFMKRFGLGPCGKWKYENIYITRQFVQFDERERHQORPCV 266
Db 181 LSSLGOKLLAFAGLSVLVLTGTLFMKRFGLGPCGKWKYENIYITRQFVQFDERERHQORPCV 240
QY 267 LPLNKEERKYYIIPFTWTPKRNKGLFELPILHLICWLVFAADVLLVRLPFSVK 326
Db 241 LPLNKEERKYYIIPFTWTPKRNKGLFELPILHLICWLVFAADVLLVRLPFSVK 300
QY 327 QFQSLPGFVHLKLGKGGTQDIIHDSSEFNISVPEPNCIPKPKFLLSETWVPLSVI 383
Db 301 QFQSLPGFVHLKLGKGGTQDIIHDSSEFNISVPEPNCIPKPKFLLSETWVPLSVI 350

RESULT 2

US-09-789-561-142
; Sequence 142, Application US/09789561
; Patent No. US20020064818A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 52 Human secreted proteins
; FILE REFERENCE: PZ043P1
; CURRENT APPLICATION NUMBER: US/09/789,561
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: PCT/US00/24008
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,317
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/152,315
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 142
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-789-561-142

Query Match 53.7%; Score 1314; DB 10; Length 291;
Best Local Similarity 92.5%; Pred. No. 4.9e-115;
Matches 259; Conservative 4; Mismatches 5; Indels 12; Gaps 3;
QY 27 MDFIQHLGVCCLVALLISVGLLSVAACWFLPSSIITAAASWIIITCVLLCCSKHARCFILLVF 86
Db 1 MDFIQHLGVCCLVALLISVGLLSVAACWFLPSSIITAAASWIIITCVLLCCSKHARCFILLVF 60
QY 87 LSCGLREGNALTAACTGIVILGHVENIHNFKGLLDGMTCNLRKSFSIHFPLLKKYIE 146
Db 61 LSCGLREGNALTAACTGIVILGHVENIHNFKGLLDGMTCNLRKSFSIHFPLLKKYIE 120
QY 147 AIQWIYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 206
Db 121 AIQWIYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 180
QY 207 LSSLGOKLLAFAGLSVLVLTGTLFMKRFGLGPCGKWKYENIYITRQFVQFDERERHQORPCV 266
Db 181 LSSLGOKLLAFAGLSVLVLTGTLFMKRFGLGPCGKWKYENIYITRQFVQFDERERHQORPCV 240
QY 267 LPLNKEERKYYIIPFTWTPKRNKGLFELPILHLICIILP-LIYLCL 306
Db 241 LPLNKEERK-----NKLKILSM-IILP-LIYLCL 268

RESULT 3

US-09-789-561-140
; Sequence 140, Application US/09789561
; Patent No. US20020064818A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 52 Human secreted proteins
; FILE REFERENCE: PZ043P1
; CURRENT APPLICATION NUMBER: US/09/789,561

CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: PCT/US00/24008
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,317
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/152,315
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 140
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-789-561-140

Query Match 53.5%; Score 1309; DB 10; Length 257;
Best Local Similarity 99.2%; Pred. No. 1.2e-114;
Matches 250; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 27 MDFIQHLGVCCLVALLISVGLLSVAACWFLPSSIITAAASWIIITCVLLCCSKHARCFILLVF 86
Db 1 MDFIQHLGVCCLVALLISVGLLSVAACWFLPSSIITAAASWIIITCVLLCCSKHARCFILLVF 60
QY 87 LSCGLREGNALTAACTGIVILGHVENIHNFKGLLDGMTCNLRKSFSIHFPLLKKYIE 146
Db 61 LSCGLREGNALTAACTGIVILGHVENIHNFKGLLDGMTCNLRKSFSIHFPLLKKYIE 120
QY 147 AIQWIYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 206
Db 121 AIQWIYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 180
QY 207 LSSLGOKLLAFAGLSVLVLTGTLFMKRFGLGPCGKWKYENIYITRQFVQFDERERHQORPCV 266
Db 181 LSSLGOKLLAFAGLSVLVLTGTLFMKRFGLGPCGKWKYENIYITRQFVQFDERERHQORPCV 240
QY 267 LPLNKEERKYY 278
Db 241 LPLNKEERKFI 252

RESULT 4

US-09-789-561-141
; Sequence 141, Application US/09789561
; Patent No. US20020064818A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 52 Human secreted proteins
; FILE REFERENCE: PZ043P1
; CURRENT APPLICATION NUMBER: US/09/789,561
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: PCT/US00/24008
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,317
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/152,315
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 141
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-789-561-141

Query Match 53.5%; Score 1309; DB 10; Length 257;
Best Local Similarity 99.2%; Pred. No. 1.2e-114;
Matches 250; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 27 MDFIQHLGVCCLVALLISVGLLSVAACWFLPSSIITAAASWIIITCVLLCCSKHARCFILLVF 86
Db 1 MDFIQHLGVCCLVALLISVGLLSVAACWFLPSSIITAAASWIIITCVLLCCSKHARCFILLVF 60
QY 87 LSCGLREGNALTAACTGIVILGHVENIHNFKGLLDGMTCNLRKSFSIHFPLLKKYIE 146

Db 61 LSCGLREGNALIAAGTGIVILGHVENIFHNFKGLDGMTNLRKSFHSIHPLKKYIE 120
Qy 147 AIQWYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 206
Db 121 AIQWYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 180
Qy 207 LSSIGQKLLAFAGLSVLVLTGTLFMKRFPGCGWKYENIYITROFQVQDERERHQPCV 266
Db 181 LSSIGQKLLAFAGLSVLVLTGTLFMKRFPGCGWKYENIYITROFQVQDERERHQPCV 240
Qy 267 LPLNKEERRKVV 278
Db 241 LPLNKEERRKFI 252

RESULT 5

US-09-798-889-55

; Sequence 55, Application US/09798889
; Publication No. US20030004324A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 31 Human secreted proteins
; FILE REFERENCE: P2026P1
; CURRENT APPLICATION NUMBER: US/09/798,889
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/393,022
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/077,714
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/077,686
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/077,687
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/077,696
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (322)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (345)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (374)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (485)
; OTHER INFORMATION: Xaa equals stop translation

US-09-798-889-55

Query Match 7.0%; Score 171; DB 9; Length 485;
Best Local Similarity 20.6%; Pred. No. 1.2e-07;
Matches 94; Conservative 63; Mismatches 167; Indels 132; Gaps 18;

Qy 35 VCCLVALLSVGLLSVAACWFLPSIIIAAASWIIITCVLLCCSKHARCFILLVFLSGLREG 94
Db 5 VCGLLVFLSLGLU-----PPPV-----RCLFALSVPITLGMBOG 36
Qy 95 RNALIAAGTGIVILGHVENIFHNFKGLDGMTNLRKSFHSIHPLKKYIEAIQWYGL 154
Db 37 RRLLSYSTALIAVAVENLVANVCAAGVLCVTEGS-----LESLLNTTHQLHAA 88
Qy 155 ATPLSVFDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEVLSLIG--Q 212
Db 89 SRALGPTCQAGSRGLT-----FEAQDNQS-----APYLHMLTIVTQVLEDFSGLE 133

Qy 213 KLLAFAGLSVLVLTGTLFMKRFPGCGW-----KYENIYITROFVQ--FDERERHQ 261
Db 134 SLARAALGQTRVVITGLFMLGLLEVSAWYLHCYLTDLRFDNIYATQTLORLAQAQATHL 193
Qy 262 QRP-----CVLPNKERRRKYVIIPTFWPTPKERKNLGLFPLFIHLICWVLFPAV 313
Db 194 LAPPPTWILQAAQLRLSQBELLSCLL-----RLGLLAL--LVATAVAV--AT 237
Qy 314 DYLLYRLIFSVSKQFOSLPGFEVHLKLHGSKQQTQDIHDSFNFISVFEP----NCIPKP 369
Db 238 DHVAFPLAQAQTVDAQKLPVPI-----TLTKYDVAVTVLGFIPFLFNQLAPES 287
Qy 370 KFL---LSETW-----VPLSV--ILLILVMLGLLSILMQLKILV 404
Db 288 PFLSVHSSYQWELRLTSARCPLLPARRPRAAFLXAGGLQLLAGSTVLLGEGYARRLRXAI 347
Qy 405 SASFYPSVERKRIQVILHAKLLKKRSKQ-----PLGE 435
Db 348 AASFFTAQEARIRRHUHLARLQRRHDXQGQQLPLGD 383

RESULT 6

US-10-205-823-367

; Sequence 367, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Womsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumel
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 367
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-367

Query Match 4.1%; Score 100.5; DB 9; Length 729;
Best Local Similarity 19.8%; Pred. No. 0.86;
Matches 79; Conservative 60; Mismatches 120; Indels 145; Gaps 16;

Qy 135 SIHFPLKKYIEAIOIYIGL-APLSVFDLVSNQTLAVSL-----FSPSHVLEA 184
Db 45 STAFIVNFECEFRFSY-YGKAVLIIFYFLYFLHWNEDTSTSIYHAFSSLCYFTP--ILGA 101
Qy 185 QLNDLS-KGEVLSVLY--QMATTTEVLSLIG-----QKLLAFAGLSVLVLTGTLFM 231
Db 102 ADADSWLGRFKTIYLSVLVVLGHVILKSLGALPILGQVHVHTVLSLIGLSIALGTG-- 158

```
QY 232 KRFLGPCGKYENIYITRQVQDERERHQORPCVLPNKERRKYV-----II 280
Db 159 --GIRPCVAAGG-----DOFEK-----HABERTRYFSVYLSINAGSLI 197
QY 281 PTFWPTPKERNKLG-----LFFLPILIHLCIWLFA-----AVDYLL 317
Db 198 STF-ITPMLRGDVOCEGDCYALAFVGLLWIALVVFAMGSKIYKPPPEGNIQAQVF 256
QY 318 YLIFSVSQFOS-----LPGFEVHL 338
Db 257 KCIMEFAISNRFRKNSGDIKPRHDWLDWAAEKYPQLIMDVKALTRVFLYIPLPMFWALL 316
QY 339 KLHGEKQGTODIHDSSFNISVPEPNCIPKPKELLSETWVPL-----380
Db 317 DOQSGRWTLQAIRMRNRNLGFFVLQPDQOMQVLNPLLIPIPLDFVIYRLVSKCGINFSS 376
QY 381 ---SVILLILVMLGLSSILMOLKILVSASFVSPVERKRIQYLH 421
Db 377 LRKMAVGMLIACLAFAVAARVEIKINEMAPAQPGQEVFLQVLN 420

RESULT 7
US-09-895-913A-234
; Sequence 234, Application US/09895913A
; Patent No. US20020160456A1
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020160456A1 Helicobacter Polypeptides in the
; FILE REFERENCE: 06132/043002
; CURRENT APPLICATION NUMBER: US/09/895,913A
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 08/891,227
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 234
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-895-913A-234

Query Match 3.9%; Score 96; DB 9; Length 405;
Best Local Similarity 25.3%; Pred. No. 1;
Matches 48; Conservative 31; Mismatches 67; Indels 44; Gaps 9;
QY 177 SPSHVLEAQLNDSKGEVLSVLYQWATTTEVLSLQKLLAFAGLSLVLLG-----TGLF 230
Db 9 TPCYVLESERLEKNAKILSEIVRQS-----GAKVL-----LALKGYAFWRBFGIL 53
QY 231 MKRFLGPCGKYENIYITR-QVQFDERERHQORPCVLPNKERRK-----YVIPT 282
Db 54 RQKNGCCA---SGLYEAKLAFEGFGRSHKEICVYSPAFKAEMSAIPLATSIIFS 110
QY 283 P--WPTPKER-----KNLGLFFLPILIHLCIWLFAAVDYLLYRLIFSVSQFQSLP 332
Db 111 FYQYATYKRIIDLKKNQLENLGL--SPIKMGLRINPLYSEVTPAIYNPCSKVSRIGITPS 168
QY 333 GFEVHLKLG 342
Db 169 GPEKGVKEHG 178

RESULT 8
US-09-924-340-40
; Sequence 40, Application US/09924340
; Publication No. US20030027248A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US2.REG
; CURRENT APPLICATION NUMBER: US/09/924,340
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Jpatent
; SEQ ID NO 40
; LENGTH: 263
```

```
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US2.REG
; CURRENT APPLICATION NUMBER: US/09/924,340
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Jpatent
; SEQ ID NO 40
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: 1..20
US-09-924-340-40

Query Match 3.9%; Score 95; DB 9; Length 263;
Best Local Similarity 24.7%; Pred. No. 0.74;
Matches 47; Conservative 23; Mismatches 72; Indels 48; Gaps 10;
QY 64 SWI---ITCVLLCCSKHARCIFILLVFLSCGLRGRNALIAAGTGVILGHVENIFHNPKG 120
Db 71 SWVGPKLPKWLHAALHMAFLVTW-----GLVAVFT-----FHN--- 106
QY 121 LLQDGTNLRKAKSFSIH--FPLKKYIEAIQWIVGLATPLSVDPDLVSNQTLAVSLFSP 178
Db 107 --HGRTANL-----YLSHWLGITTVDFGQWFLGFA-----VF--LLPWASWNLRLSKP 154
QY 179 SHVL--EAQLNDSKGEVLS-----VLYQWATTTEVLSLQKLLAFAGLSLVLLGTGLFM 231
Db 155 IHVFFGAAILSLSTASVISGINKLFPSLKNLTPRPHSLPSEAVFANSTGMLVVAFGLLV 214
QY 232 KRFLGPCGKW 241
Db 215 LYILLASSWK 224

RESULT 9
US-09-992-600A-40
; Sequence 40, Application US/09992600A
; Publication No. US20030027161A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US4.DIV
; CURRENT APPLICATION NUMBER: US/09/992,600A
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Jpatent
; SEQ ID NO 40
; LENGTH: 263
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: 1..20
US-09-992-600A-40

Query Match      3.9%; Score 95; DB 9; Length 263;
Best Local Similarity 24.7%; Pred. No. 0.74;
Matches 47; Conservative 23; Mismatches 72; Indels 48; Gaps 10;

Qy 64 SWI---ITCVLLCCSKHARCFILLVFLSCGLREGRNALIAAGTGIVILGHVENIFHNFKG 120
Db 71 SWVGPKLPWKLHAAHLMAFVLTVV-----GLVAVFT-----FHN--- 106

Qy 121 LLDGMTCNLRKSFSIH--FPLLKKYIEAIQWYGLATPLSVFDDLVSNQTLAVSLFSP 178
Db 107 --HGRTANL---YSLHSLWLGITTVFLFCQWFLGFA----VF--LLPWASWMLRSLKLP 154

Qy 179 SHVL---EAQLNDSKGEVLS-----VLYQMATTTEVLSLQKLLAFAGLSVLVLTGLTFM 231
Db 155 IHVFFGAAILSLSIASVISGINEKLFSLKNTTRPYHSLPSEAVFANSTGMLVWVAFGLIV 214

Qy 232 KRFLGPCGKW 241
Db 215 LYILLASSWK 224

RESULT 10
US-10-000-489-40
; Sequence 40, Application US/10000489
; Publication No. US20030092011A1
; GENERAL INFORMATION:
; APPLICANT: Benjanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US6.DIV
; CURRENT APPLICATION NUMBER: US/10/000,489
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 40
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: 1..20
US-10-000-489-40

Query Match      3.9%; Score 95; DB 9; Length 263;
Best Local Similarity 24.7%; Pred. No. 0.74;
Matches 47; Conservative 23; Mismatches 72; Indels 48; Gaps 10;

Qy 64 SWI---ITCVLLCCSKHARCFILLVFLSCGLREGRNALIAAGTGIVILGHVENIFHNFKG 120
Db 71 SWVGPKLPWKLHAAHLMAFVLTVV-----GLVAVFT-----FHN--- 106

Qy 121 LLDGMTCNLRKSFSIH--FPLLKKYIEAIQWYGLATPLSVFDDLVSNQTLAVSLFSP 178
Db 107 --HGRTANL---YSLHSLWLGITTVFLFCQWFLGFA----VF--LLPWASWMLRSLKLP 154
```

```
Qy 179 SHVL---EAQLNDSKGEVLS-----VLYQMATTTEVLSLQKLLAFAGLSVLVLTGLTFM 231
Db 155 IHVFFGAAILSLSIASVISGINEKLFSLKNTTRPYHSLPSEAVFANSTGMLVWVAFGLIV 214

Qy 232 KRFLGPCGKW 241
Db 215 LYILLASSWK 224

RESULT 11
US-10-000-986-40
; Sequence 40, Application US/10000986
; Publication No. US20030096247A1
; GENERAL INFORMATION:
; APPLICANT: Benjanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US9.DIV
; CURRENT APPLICATION NUMBER: US/10/000,986
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 40
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: 1..20
US-10-000-986-40

Query Match      3.9%; Score 95; DB 9; Length 263;
Best Local Similarity 24.7%; Pred. No. 0.74;
Matches 47; Conservative 23; Mismatches 72; Indels 48; Gaps 10;

Qy 64 SWI---ITCVLLCCSKHARCFILLVFLSCGLREGRNALIAAGTGIVILGHVENIFHNFKG 120
Db 71 SWVGPKLPWKLHAAHLMAFVLTVV-----GLVAVFT-----FHN--- 106

Qy 121 LLDGMTCNLRKSFSIH--FPLLKKYIEAIQWYGLATPLSVFDDLVSNQTLAVSLFSP 178
Db 107 --HGRTANL---YSLHSLWLGITTVFLFCQWFLGFA----VF--LLPWASWMLRSLKLP 154

Qy 179 SHVL---EAQLNDSKGEVLS-----VLYQMATTTEVLSLQKLLAFAGLSVLVLTGLTFM 231
Db 155 IHVFFGAAILSLSIASVISGINEKLFSLKNTTRPYHSLPSEAVFANSTGMLVWVAFGLIV 214

Qy 232 KRFLGPCGKW 241
Db 215 LYILLASSWK 224

RESULT 12
US-10-180-946-2
; Sequence 2, Application US/10180946
; Publication No. US20030093825A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Keith D.
; TITLE OF INVENTION: PUTATIVE NEUROTRANSMITTER RECEPTOR (PNR)
; TITLE OF INVENTION: DISRUPTIONS, COMPOSITIONS AND METHODS RELATING THERETO
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2003, 13:24:17 ; Search time 111 Seconds
(without alignments)
1084.783 Million cell updates/sec

Title: US-09-713-098-2

Perfect score: 2448

Sequence: 1 MGWISGTDIFLSLWEIYVS.....LPVLKMRKKQMDASADKS 470

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1160474 seqs, 256193413 residues

Total number of hits satisfying chosen parameters: 1160474

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New.*

1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1314	53.7	291	5	US-09-950-083B-3966 Sequence 3966, Ap
2	1309	53.5	257	5	US-09-950-083B-3965 Sequence 3965, Ap
3	754.5	30.8	148	6	US-10-264-237-2360 Sequence 2360, Ap
4	233.5	9.5	639	6	US-10-094-749-3184 Sequence 3184, Ap
5	207	8.5	706	6	US-10-094-749-3082 Sequence 3082, Ap
6	115.5	4.7	526	1	PCT-US02-36123-1586 Sequence 1586, Ap
7	115.5	4.7	553	1	PCT-US02-36123-1588 Sequence 1588, Ap
8	115.5	4.6	559	1	PCT-US02-36123-1590 Sequence 1590, Ap
9	112	4.6	1507	6	US-10-437-963-200793 Sequence 200793, Ap
10	105.5	4.3	464	6	US-10-282-122A-72462 Sequence 72462, A
11	104.5	4.3	328	6	US-10-282-122A-73382 Sequence 73382, A
12	104.5	4.3	496	7	US-60-419-463-39 Sequence 39, Appl
13	104	4.2	2275	7	US-60-452-680-23634 Sequence 23634, A
14	103	4.2	591	1	PCT-US03-06367-42 Sequence 42, Appl
15	102.5	4.2	1494	1	PCT-US02-22217-17 Sequence 17, Appl
16	102.5	4.2	1494	1	PCT-US03-01343-17 Sequence 17, Appl
17	102.5	4.2	1494	1	PCT-US03-01343A-17 Sequence 17, Appl
18	102.5	4.2	1494	6	US-10-345-072-17 Sequence 17, Appl
19	101.5	4.1	738	5	US-09-949-016-11165 Sequence 11165, A
20	100.5	4.1	729	5	US-09-949-016-6988 Sequence 6988, Ap
21	100.5	4.1	729	6	US-10-219-051B-7260 Sequence 7260, Ap
22	100.5	4.1	729	6	US-10-219-051B-10505 Sequence 10505, A
23	100.5	4.1	729	7	US-60-457-901-8 Sequence 8, Appl
24	100	4.1	611	6	US-10-369-493-10288 Sequence 10288, A
25	99	4.0	242	6	US-10-305-278-298 Sequence 298, Ap
26	99	4.0	242	6	US-10-094-749-2757 Sequence 2757, Ap

27 99 4.0 485 5 US-09-949-016-10455 Sequence 10455, A
28 99 4.0 530 5 US-09-949-016-8316 Sequence 8316, Ap
29 99 4.0 879 5 US-09-724-676-66581 Sequence 66581, A
30 99 4.0 879 5 US-09-724-676A-66581 Sequence 66577, A
31 99 4.0 924 5 US-09-724-676-66577 Sequence 66577, A
32 99 4.0 924 5 US-09-724-676-66577 Sequence 66578, A
33 99 4.0 924 5 US-09-724-676-66579 Sequence 66579, A
34 99 4.0 924 5 US-09-724-676-66580 Sequence 66580, A
35 99 4.0 924 5 US-09-724-676A-66577 Sequence 66577, A
36 99 4.0 924 5 US-09-724-676A-66578 Sequence 66578, A
37 99 4.0 924 5 US-09-724-676A-66579 Sequence 66579, A
38 99 4.0 924 5 US-09-724-676A-66580 Sequence 66580, A
39 99 4.0 997 6 US-10-348-119-238 Sequence 238, App
40 99 4.0 1042 5 US-09-724-676-66584 Sequence 66584, A
41 99 4.0 1042 5 US-09-724-676-66585 Sequence 66585, A
42 99 4.0 1042 5 US-09-724-676-66586 Sequence 66586, A
43 99 4.0 1042 5 US-09-724-676A-66584 Sequence 66584, A
44 99 4.0 1042 5 US-09-724-676A-66585 Sequence 66585, A
45 99 4.0 1042 5 US-09-724-676A-66586 Sequence 66586, A

ALIGNMENTS

RESULT 1

US-09-950-083B-3966
; Sequence 3966, Application US/09950083B
; GENERAL INFORMATION:

APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS805
; CURRENT APPLICATION NUMBER: US/09/950,083B
; CURRENT FILING DATE: 2003-05-07
; PRIOR APPLICATION NUMBER: 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US01/11988
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US00/06043
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06012
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06058
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06044
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06059
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06042
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06014
; PRIOR FILING DATE: 2000-03-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13046
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3966
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-950-083B-3966

Query Match 53.7%; Score 1314; DB 5; Length 291;
Best Local Similarity 92.5%; Pred. No. 6.5e-118;
Matches 259; Conservative 4; Mismatches 5; Indels 12; Gaps 3;
QY 27 MDFOHLGVCCVALISVGLLSVAACWFLPSIIIAAASWIIITCVLLCCSKHARCFILLVF 86
Db 1 MDFOHLGVCCVALISVGLLSVAACWFLPSIIIAAASWIIITCVLLCCSKHARCFILLVF 60
QY 87 LSCGLREGNALLAAGTGIVILGHVENIFHNFKGLDGMTCNLRKSFSTHFFLLKKYIE 146
Db 61 LSCGLREGNALLAAGTGIVILGHVENIFHNFKGLDGMTCNLRKSFSTHFFLLKKYIE 120

```

147 AIQWYGLATPLSVFDDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLQMATTTTEV 206
121 AIQWYGLATPLSVFDDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLQMATTTTEV 180
207 LSSLGQKLLAFAGLSVLVLTGTLFMKRLFGCGWKYENIYITRQVQFDERERHQRPCV 266
181 LSSLGQKLLAFAGLSVLVLTGTLFMKRLFGCGWKYENIYITRQVQFDERERHQRPCV 240
267 LPLNKERRKVIITPTWPTPKERRKNLGLFFLPLILHLCL 306
241 LPLNKERRK-----NKLKILISM-TLP-LIYLCL 268

```

RESULT 2

INVENTOR: Z
 US-09-950-083B-3965
 ; Sequence 3965, Application US/09950083B
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen, et. al
 ; TITLE OF INVENTION: Human Secreted Proteins
 ; FILE REFERENCE: PS805

Query Match	53.5%	Score 1309;	DB 5;	Length 257;
Best Local Similarity	99.2%;	Pred. No. 1.7e-117;		
Matches 250;	Conservative	2;	Mismatches 0;	Indels 0;
Gaps				

Qy		87	LSCGLREGRNALIAAGTGTIVILGHVENI	IHNFKGILLDGMTCNLRKSF	IHFPLLKYYIE	146
Dd		61	LSCGLREGRNALIAAGTGTIVILGHVENI	IHNFKGILLDGMTCNLRKSF	IHFPLLKYYIE	120

Qy	147	AIQWYIGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQIANDSKGEVLSVLVYQMATTEV	206
Db	121	AIQWYIGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQIANDSKGEVLSVLVYQMATTEV	180

Qy	207	LSLSGOKLLAFAGLSLVLLGTGLFMKRFGLPCGWKYENIITROFVQFDERERHQORPCV	2666
Db	181	LSLSGOKLLAFAGLSLVLLGTGLFMKRFGLPCGWKYENIITROFVQFDERERHQORPCV	240

QY 267 LPLNKEERRKYV 278

Db 241 LPLNKEERRKEI 252

RESULT 3

```

US-10-264-237-2360
; Sequence 2360, Application US/10264237
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn ver. 3.1
; SEQ ID NO 2360
; LENGTH: 148

```

Query Match 30.8%; Score 754.5; DB 6; Length 148;
Best Local Similarity 91.5%; Pred. NO. 1.8e-64;
Matches 140; Conservative 2; Mismatches 4; Indels 7; Gaps 1

291 KNGLFFLPILIHLCIWLFAAVDYLRLIFSVSKQFQSLPGFEVHLKLGKQGTDI 350
 61 KNGLFFLPILIHLCIWLXAAVDYLRLIFSVSKQFQSLPGFEVHLKLGKQGTDI 120

y	351	I	H	D	S	S	F	N	I	S	V	F	E	P	N	C	I	P	K	P	F	L	S	E	T	W	V	P	L	S	V	383
b	121	I	H	D	S	S	F	N	I	S	V	F	E	P	N	C	I	P	K	P	-----	W	O	A	L	K	L	146				

RESULT 4

US-10-094-749-3184
; Sequence 3184, Application US/10094749
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO

```
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3184
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-3184

Query Match          9.5%; Score 233.5; DB 6; Length 639;
Best Local Similarity 20.8%; Pred. No. 1.8e-13;
Matches 127; Conservative 88; Mismatches 194; Indels 201; Gaps 24;

QY 1 MGIWTSQTDIFUS--LMEIYVSPSPGWMDFQHGLGVCCCLVALISVGLLSVAACWFLPSI 58
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
27 VGGFTLGLSLATAYAGLELLVEGHP--WGCLV---GTLTLAFLSLG-----69
QY 59 TAAASWIIITCVLLCCSKHARCFILLVFLSCGLRGNALIAAGTGIVILGHVENIFHNF 118
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
70 -----MGFSROVRATVLLLLPQAFSRQGTLLVAAAFGLVQGPCCANTLRF 116
QY 119 KGLLDGMCNL-----RAKFSIHPEPL-----LKKYIE 146
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
117 TRASEAVACGAEALNQTAEVLQRAQ-----PLVSALNKIKAIARKTEVADRVRKFFR 171
QY 147 AI-----OWIY-----152
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
172 SIMDGVKHITARLRNVWQWLLHIGDVCNSELGNPYLKARVFDADKSCMWVIPAAYHLC 231
QY 153 -----GLATPLSVFDDLVSW-----NQTALVSL-----FSPSHVL 182
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
232 YVLMPPKALCALGLASLVQFCVPIPKYIQPFLRQTGTGTPVQLNVRVQEFENMTATHF 291
QY 183 EAQINDSKEVLSVLQWATTEVLSLQKL-----LAFAGLSVLVLTGTLFMKR-PL 235
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
292 SVDLNASRS-----LSOVA--MDLHEAVSMKLRVREALALMGFTTPLLVLVLLQALFY 344
QY 236 GPC--GW-KYENIYITROPVQDERERHQRCVPLNKEERRKVIIPTFWPTPERKN 292
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
345 RYCYLNWDHYDNIYITSRFLRMEAVRSTAGTLVPLSAHEARRVIPPQSI FLSQWERFF 404
QY 293 LGLFELPILHLCIWFLEAAVDYLLYRLIFSVSKQFQS--LPGFEVHLKLHGEKQ--TOD 349
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
405 YLETFNLIRHLLVLFLVFLDYAVFWVLDLHQLQGEIVARSPVLSLTVEGTGYAGN 464
QY 350 IHD--SSP-----NISVFENCIPKPKFLSETWVPLSVILLILVMLGLSSILMQLK 401
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
465 IYRDLVSAFDVLQQGNISLRSCLLRPSEPDPSTGYIVIGVMYGLCFFITLFGSVVRLR 524
QY 402 ILVASFYSVERKRIQVILHAKLKKRSKQPLG---EVKRL-----SLYLTKI-----447
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
525 RVICASYYSRQERISLYNVLSSRTNLLAALHRSVRRRAADQGHRSFAFLVLSRCPFC 584
QY 448 -----HFWL 451
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
585 LQPFVSHFWL 594

RESULT 5
US-10-094-749-3082
; Sequence 3082, Application US/10094749
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
```

```
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094, 749
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3082
; LENGTH: 706
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-3082

Query Match          8.5%; Score 207; DB 6; Length 706;
Best Local Similarity 19.5%; Pred. No. 7.4e-11;
Matches 123; Conservative 86; Mismatches 172; Indels 250; Gaps 23;

QY 6 SGTDPLSLWEIYVSPSPGWMDFQHGLGVCCCLVALISVGLLSVAACWFLPSIAAASW 65
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
56 AGGLAIGFOLLVNP-----MNIYBEQIMPLYSLVGLG-----AMGW 94
QY 66 IITCVLLCCSKHARCFILLVFLSCGLRGNALIAAGTGIVILGHVENIFHNFKGLDGM 125
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
95 -----GTSPIRHCASLLLVPMKLGKGRFLFVGLAALAIYGPVANLRHNNVITSL 147
QY 126 TC-----NLRAKFSIHFPPLKKYIEATQWYIGLATPLSVFDDLVSMNQTAV---SL 175
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
148 GCTVELQINTRA-----AMRISTAPLRAFMFKDLLSSKELLRAETRN 190
QY 176 FSPSHVLEAQLNDSKEV-----LSV--LYOMATTTEVLS 208
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
191 SATFEDLDAQVNSGTGTPEDTMDSGETAQGREARQAPASRLHLSSTOKMYELTKLRCSY 250
QY 209 SLGOKLLA-----FAGLSVL-----224
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
251 VVQQAILSCRRWFDRKHQCMKHIWVPLLTLLCLPMKFKFCGIAKMEVWCNRPVE 310
QY 225 -----L 225
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
311 GNFGQTYDSLNSIRGLDGEFSANIDFKEQAGVLGNTSWERVSTEVDDYVVRQEARL 370
QY 226 GTGLPMKRLGPGCW-----KYENIYITROPVQDERERHQRCVPL 267
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
371 ENAIGLHLVLLSCTPLLVLHASFSYMSYNDHIFDNIIYISTVFCDDRRKKLGRKTL 430
QY 268 PLNKEERRKVIIPTFWPT--PKERNLG---LFFLP-----LIHLC--IWLFAAVDYLL 317
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
431 PLRKAE-EXTVIFPC-KPTIQASEMSNVVRELLETLPILLVVLVCLGDWALYSIFDTIR 488
QY 318 YRLIFSVMKQFOSLPGFEVHLKLHGEK-----QGTODIHDHDSFNISFEPN---CIPK 368
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
489 HHSFLQYS--FRSSHKLEV--KVGDSLLARLURKTIGALNTSS--ETVMESNMNMCPLQ 542
QY 369 PKFLLSETWVPLSVILLILVMLGLSSILMQLKILVSASFYSVERKRIQVILHAKLKKR 428
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
543 PVGLDARAYWRAAVPIGLLVCLLQAFGVRLRRTVIAAFYFKREKRIILFLYNDLLKKR 602
QY 429 S-----KQPLGEVKRR 439
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
603 AAFTKLRRRAAILRRERQKAPRHPPLADILHR 633
```



```
RESULT 6
PCT-US02-36123-1586
; Sequence 1586, Application PC/TUS0236123
; GENERAL INFORMATION:
; APPLICANT: American Cyanamid Company, and Fletcher, Leah D., McMichael, John C.,
; APPLICANT: Russell, David P., and Zagursky, Robert J.
; TITLE OF INVENTION: Alloiooccus otitidis Open Reading Frames (ORFs) Encoding Polypep
; FILE OF INVENTION: Antigens, Immunogenic Compositions and Uses Thereof
; FILE REFERENCE: Application 1
; CURRENT APPLICATION NUMBER: PCT/US02/36123
; CURRENT FILING DATE: 2003-01-02
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1586
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Alloiooccus otitidis
PCT-US02-36123-1586

Query Match          4.7%; Score 115.5; DB 1; Length 526;
Best Local Similarity 22.1%; Pred. No. 0.031;
Matches 79; Conservative 49; Mismatches 103; Indels 127; Gaps 16;

QY 84 LVFLSCGLRE-----GRNALIAAGTGIVILGHVENIFHNPKGLLDGMCNLR----- 130
DB 89 MLFLSTDLTEAGITTDYLGTSGLIAAFVSAVI---VVNTYYF--IKKNITINLPNEVPG 143
QY 131 --AKSFHIFLLKKYIEAIOIYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLND 188
DB 144 NVASAFEDIMPVSII--LIYVDLVTRYFVGNNAQF---VIELFQP----- 187
QY 189 SKGEVLSVLYQMATTTEVLSLQKLLAFA-----GLSLVLLGTGLFMRKRLGPC 238
DB 188 -----LFQVAD-----SYVGLIIIAFAVSFFWFIGHGPSVLPVATM----- 227
QY 239 GWKYNIIYITROFQFDERERHQORPC-----VLP-----LNKEERKXY 278
DB 228 ---YQNLNQLIQAEGAFHTLTPTWNSFVITMGCTGATFVVPFIFMWLARSKRKAV 284
QY 279 IIPTFWPTPKERNKGLFLPLILHLCIWVLF--AVDYLLYRL---IFSVSQKQFOSLP 332
DB 285 GKAAFIPTSGVNEPILFGGPLIINPIFFVFPVLPVNTVILYKLVDFLFSMNA----- 338
QY 333 GFEVHLKLHGEKQGTODIHDSSFNISVPEPCIPKPKLLSET--WVPLSVILLILVM 389
DB 339 -----LSLYLPMTTPGPGLVMTSGFAPLSFILAILIL 371

RESULT 7
PCT-US02-36123-1588
; Sequence 1588, Application PC/TUS0236123
; GENERAL INFORMATION:
; APPLICANT: American Cyanamid Company, and Fletcher, Leah D., McMichael, John C.,
; APPLICANT: Russell, David P., and Zagursky, Robert J.
; TITLE OF INVENTION: Alloiooccus otitidis Open Reading Frames (ORFs) Encoding Polypep
; FILE OF INVENTION: Antigens, Immunogenic Compositions and Uses Thereof
; FILE REFERENCE: Application 1
; CURRENT APPLICATION NUMBER: PCT/US02/36123
; CURRENT FILING DATE: 2003-01-02
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1588
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Alloiooccus otitidis
PCT-US02-36123-1588

Query Match          4.7%; Score 115.5; DB 1; Length 553;
Best Local Similarity 22.1%; Pred. No. 0.034;
Matches 79; Conservative 49; Mismatches 103; Indels 127; Gaps 16;

QY 84 LVFLSCGLRE-----GRNALIAAGTGIVILGHVENIFHNPKGLLDGMCNLR----- 130
DB 89 MLFLSTDLTEAGITTDYLGTSGLIAAFVSAVI---VVNTYYF--IKKNITINLPNEVPG 143
QY 131 --AKSFHIFLLKKYIEAIOIYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLND 188
DB 144 NVASAFEDIMPVSII--LIYVDLVTRYFVGNNAQF---VIELFQP----- 187
QY 189 SKGEVLSVLYQMATTTEVLSLQKLLAFA-----GLSLVLLGTGLFMRKRLGPC 238
DB 188 -----LFQVAD-----SYVGLIIIAFAVSFFWFIGHGPSVLPVATM----- 227
QY 239 GWKYNIIYITROFQFDERERHQORPC-----VLP-----LNKEERKXY 278
DB 228 ---YQNLNQLIQAEGAFHTLTPTWNSFVITMGCTGATFVVPFIFMWLARSKRKAV 284
QY 279 IIPTFWPTPKERNKGLFLPLILHLCIWVLF--AVDYLLYRL---IFSVSQKQFOSLP 332
DB 285 GKAAFIPTSGVNEPILFGGPLIINPIFFVFPVLPVNTVILYKLVDFLFSMNA----- 338
QY 333 GFEVHLKLHGEKQGTODIHDSSFNISVPEPCIPKPKLLSET--WVPLSVILLILVM 389
DB 339 -----LSLYLPMTTPGPGLVMTSGFAPLSFILAILIL 371

RESULT 8
PCT-US02-36123-1590
; Sequence 1590, Application PC/TUS0236123
; GENERAL INFORMATION:
; APPLICANT: American Cyanamid Company, and Fletcher, Leah D., McMichael, John C.,
; APPLICANT: Russell, David P., and Zagursky, Robert J.
; TITLE OF INVENTION: Alloiooccus otitidis Open Reading Frames (ORFs) Encoding Polypep
; FILE OF INVENTION: Antigens, Immunogenic Compositions and Uses Thereof
; FILE REFERENCE: Application 1
; CURRENT APPLICATION NUMBER: PCT/US02/36123
; CURRENT FILING DATE: 2003-01-02
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1590
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Alloiooccus otitidis
PCT-US02-36123-1590

Query Match          4.7%; Score 115.5; DB 1; Length 559;
Best Local Similarity 22.1%; Pred. No. 0.034;
Matches 79; Conservative 49; Mismatches 103; Indels 127; Gaps 16;

QY 84 LVFLSCGLRE-----GRNALIAAGTGIVILGHVENIFHNPKGLLDGMCNLR----- 130
DB 122 MLFLSTDLTEAGITTDYLGTSGLIAAFVSAVI---VVNTYYF--IKKNITINLPNEVPG 176
QY 131 --AKSFHIFLLKKYIEAIOIYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLND 188
DB 177 NVASAFEDIMPVSII--LIYVDLVTRYFVGNNAQF---VIELFQP----- 220
QY 189 SKGEVLSVLYQMATTTEVLSLQKLLAFA-----GLSLVLLGTGLFMRKRLGPC 238
DB 221 -----LFQVAD-----SYVGLIIIAFAVSFFWFIGHGPSVLPVATM----- 260
QY 239 GWKYNIIYITROFQFDERERHQORPC-----VLP-----LNKEERKXY 278
DB 261 ---YQNLNQLIQAEGAFHTLTPTWNSFVITMGCTGATFVVPFIFMWLARSKRKAV 317
QY 279 IIPTFWPTPKERNKGLFLPLILHLCIWVLF--AVDYLLYRL---IFSVSQKQFOSLP 332
DB 318 GKAAFIPTSGVNEPILFGGPLIINPIFFVFPVLPVNTVILYKLVDFLFSMNA----- 371
QY 333 GFEVHLKLHGEKQGTODIHDSSFNISVPEPCIPKPKLLSET--WVPLSVILLILVM 389
DB 372 -----LSLYLPMTTPGPGLVMTSGFAPLSFILAILIL 404

RESULT 9
```

```
DB 116 MLFLSTDLTEAGITTDYLGTSGLIAAFVSAVI---VVNTYYF--IKKNITINLPNEVPG 170
QY 131 --AKSFHIFLLKKYIEAIOIYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLND 188
DB 171 NVASAFEDIMPVSII--LIYVDLVTRYFVGNNAQF---VIELFQP----- 214
QY 189 SKGEVLSVLYQMATTTEVLSLQKLLAFA-----GLSLVLLGTGLFMRKRLGPC 238
DB 215 -----LFQVAD-----SYVGLIIIAFAVSFFWFIGHGPSVLPVATM----- 254
QY 239 GWKYNIIYITROFQFDERERHQORPC-----VLP-----LNKEERKXY 278
DB 255 ---YQNLNQLIQAEGAFHTLTPTWNSFVITMGCTGATFVVPFIFMWLARSKRKAV 311
QY 279 IIPTFWPTPKERNKGLFLPLILHLCIWVLF--AVDYLLYRL---IFSVSQKQFOSLP 332
DB 312 GKAAFIPTSGVNEPILFGGPLIINPIFFVFPVLPVNTVILYKLVDFLFSMNA----- 365
QY 333 GFEVHLKLHGEKQGTODIHDSSFNISVPEPCIPKPKLLSET--WVPLSVILLILVM 389
DB 366 -----LSLYLPMTTPGPGLVMTSGFAPLSFILAILIL 398
```

US-10-437963-200793

; Sequence 200793, Application US/10437963

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with

; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 200793

; LENGTH: 1507

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_96228C.1.pep

US-10-437-963-200793

Query Match 4.6%; Score 112; DB 6; Length 1507;

Best Local Similarity 20.4%; Pred. No. 0.29;

Matches 103; Conservative 66; Mismatches 179; Indels 158; Gaps 21;

Qy 54 PLPSITIAAASWITCVLLCCSKNRCFILLVFLSCGLREGNALLAAGTGIVILGHVEN 113

Db 931 FRPICLLNVSPKIIITKVLNNRIYHVMCYIIISKNQSAFL---KNRYIL---EGVILHEILN 985

Qy 114 IFHNFK--GLDGMTCNLRKSFSTHFPLLKKYIEAIOIY----- 152

Db 986 SLHKKQSGIL-----FKVD---LEKAYDKNVVVFHRLMVLKSGPPDQWIM 1031

Qy 153 ----GLATPLSVFDDLVSNQT-----LAVLSFSPSHVLEAQLNDSKGEVLSVLYQ 199

Db 1032 KVIKGGKVAKVNDIEIGLFTQHKGLRQGDPLSPLEFLNLA-----AEALTLLVQ 1080

Qy 200 MATTEVLSSIGQ-----KLLAFAGLSVLVLTGTFMCRFLGPGCWKYE-----NIY 246

Db 1081 RAEBSHLEGLGNNEENKISILQVADDTIFLLPDNLEARNLKFTILCFEQLSDLKINFH 1140

Qy 247 ITROPVQDERERH-----QQRPCVLPLNKEER-----RKYVI 279

Db 1141 KSEVFCIGEARDKVFLSQLFTCNIGDLPMLRPTGVKERMDFFRRLRLWQEDQGIRKYHL 1200

Qy 280 IPTFWP---TPKERNKGLFLPI-----LIHLCIWVLPAAVDYLLYRLIFSVMKQFSLP 332

Db 1201 VQ--WPVICTRDOEGLGVLDLDMNKALLGKWIW----- 1233

Qy 333 GFEVHLKHLGKQGTQDIHDSRNISVFENCIPKPKFLLSETWVPLSVIL-----LIL 387

Db 1234 -----KLENEGDWQEVLRDKYLKCKTMS---IQPKPGDSHFQWGLMEVKERYNFCT 1284

Qy 388 VMGLLSILMQLKILVS-----ASFYPSVERKRIQYLHAKLLKKRSKPLGEVKRRLSY 443

Db 1285 KKVGDGRSIMPWEDSWLGGRLAHLFPTLYNITLS-KHVSLEIKNK-GWEVIKFRRLSY 1342

Qy 444 LTKIHFVPLVMIRKKQMDMASADK 469

Db 1343 GSKLRDWNKIKLSVEGIOVEGVDRDK 1368

RESULT 10

US-10-282-122A-72462

; Sequence 72462, Application US/10282122A

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Cart, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELIPIRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 72462

; LENGTH: 464

; TYPE: PRT

; ORGANISM: Streptococcus mutans

US-10-282-122A-72462

Query Match 4.3%; Score 105.5; DB 6; Length 464;

Best Local Similarity 19.3%; Pred. No. 0.24;

Matches 91; Conservative 72; Mismatches 160; Indels 149; Gaps 22;

Qy 33 LGVCCVLVALISGLSVAAACWFLPSIIAAAA-----SWII---TCVLCC-----SKHA- 78

Db 13 LGI--FIVMLDTTMMIT----LPAIQKGLGVKLDQLSWNTINYTIIFASCTIPLSKIAD 66

Qy 79 ----RCFI--LLVFLSCGLREGNALLAAGTGIVILGHVENIFHNFKGLDGMTCNLRA 131

Db 67 IVGKRLFLVIGLLPFGISLLSG----LANGFSQLILGRIISSFGAAILLPVNSLGISS 122

Qy 132 KSFSIHFPLKKYIEAIOIYIGLATPL-----SVFDDLVSNQTLAVLSFSPSHVLEAQL 186

Db 123 WEVKDRF-----KIVAAALGLMGQGAARAIGPTLGGILDTFSWHWIFINL--PIIITATCL 176

Qy 187 NDSKGEVSVLVYOMATTTTEVLSLQKLLAFAGLSVLVLTGTFMCRFLGPGCWKKNYIY 246

Db 177 ----MILSYHPKSEKIESKIDFAGSFISMLGLFLVTL--GLIKIDWAGDWRTLGCL 229

Qy 247 IT--RQFVQFDERERHQPCVLPINKERRKY-----VIPTFWPT 286

Db 230 ITFLLSLFAFIILEKHSKNP-MINLNLFKIREFTASALVALLAQQFYIGIVITLPTFFTT 288

Qy 287 PKERNL--GLFFLP--ILIHLC-----IWVLFPAVDYLLYRLIFSVS- 325

Db 289 IOGKTDLAALLPMLSVVFFICGGGLSVINQLGPRLIVFVGLTAI--LLSYLLIVSINP 347

Qy 326 -----KQPSLFGGEVHL 338

Db 348 NKVMAMALTTILIGIFGFIAGIPVNVVLAASTLQGELLTASQSIVGVVRQIGSVLGVTVFI 407

;; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;; FILE REFERENCE: CL001450
;; CURRENT APPLICATION NUMBER: US/60/452,680
;; CURRENT FILING DATE: 2003-03-07
;; NUMBER OF SEQ ID NOS: 116213
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 23634
;; LENGTH: 2275
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-60-452-680-23634

Query Match 4.2%; Score 104; DB 7; Length 2275;
Best Local Similarity 20.3%; Pred. No. 3;
Matches 107; Conservative 77; Mismatches 160; Indels 184; Gaps 28;
QY 6 SGTDFLSLWELVSPRPGWMDFIQHL---GVCC---LVALISVGLLSVAACWFLPSI 58
Db 1492 TGFDFVSDLE-HISPHQ--MTSLQYLHAQPTICQGFPCAIR-ALHQHCACKKHPOW 1547
QY 59 IAAASWIIITCVLLCCSHARCIFILLVFLSCGLREGRNALIAAGTGVILGHVENIFHNF 118
Db 1548 IG-----LITSLPYMGK---VLQVVVSVTLQCRN-----LDNLIQY 1584
QY 119 K-----GLDGMTCNLRKAKSFIHPILKKYIEAIOIYGLA 155
Db 1585 KYETGLSDSRPLMWASIIIPDMILTLREGITA-----IIHYCLL----- 1623
QY 156 TPLSVFDDVSNQTLAVLSLSPSHVLEAQLNDSKEVLSVLYQMATTTEVLSSL----- 210
Db 1624 -----DPTQYHQ-LLVSV-DQRHLFEA-----RSGILSILHMTSSVTLLWSILHQAD 1670
QY 211 -GOKLAPAGLSVLVLTGLF-----MKRFLGPCWKYENIYITRFQVQDERERHO-- 261
Db 1671 SSEKWTIAASLTTINLGATKNLRQQLLELLGPISMHGVHFMFAIAFVWNERQNKIT 1730
QY 262 QRCVPLPUNKER-----RKVIIPTEWPT-----PKERKNLGLFPLPIL 301
Db 1731 TRTKVIPAASEQLLVLEVRISVWRAETVIQTKEVLKQPPATAKDKHLSL----- 1784
QY 302 IHLCTIWLFAADVLLYRLIFSVPK-QSLPGFEVHLKLHGEKQGTQDIHDSFENISV 360
Db 1785 -EVCMLQFF-----YAIQVRLNEFIKNPS-----LENKKDQDRLQDVTHKIVDAIGA 1832
QY 361 FEPNCI-----BPKFLLSETWVPLSVLLILVLMGLLSLMLQKLIVSAS 407
Db 1833 IAGSLEQTTWLRNLEVPKIMVDGNTNLESDEGILVHKHGLL-SVLAHLDMV--- 1888
QY 408 FYPVSVERKRIQYLHAKLLKGRSKQPLGSEVGRSLSLYTKIHFWLPVLK 455
Db 1889 FY-SDEKERV-----IPLLVNIMHVVPYLR 1913

RESULT 14

PCT-US03-06367-42
; Sequence 42, Application PC/TUS0306367
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: The Brigham and Women's Hospital, Inc.
; APPLICANT: Case Western Reserve University
; APPLICANT: Mount, David B
; APPLICANT: Romero, Michael
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF SLC26A8 and SLC11 ANION
; FILE REFERENCE: EXCHANGERS
; FILE REFERENCE: 1242/50/4 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/06367
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,828
; PRIOR FILING DATE: 2002-02-28
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 42

;; LENGTH: 591
;; TYPE: PRT
;; ORGANISM: Drosophila melanogaster
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (583)..(583)
;; OTHER INFORMATION: The 'Xaa' at location 583 stands for Gly, or Val.
PCT-US03-06367-42
Query Match 4.2%; Score 103; DB 1; Length 591;
Best Local Similarity 22.3%; Pred. No. 0.59;
Matches 80; Conservative 54; Mismatches 113; Indels 112; Gaps 20;
QY 1 MGIWTSGTDFLSLWELVSP-RSPGWMDFIOHLGVCCIVALLSVGLLSVAACWFLPSII 59
Db 213 LGIKTSG-NTFLQVWISIVNDIHNIWSPDPI--LGIVCITLLLS--LRALASCTLGPKEG 267
QY 60 AAAASWIIITCVLLCCSHARCIFILLVFLSCGLREGRNALIAAGTGVILGHVENIFHNF 119
Db 268 KTAQKLLTG-----IFWTIG--TARNALLVCGT--AGLGYW--LFVNGK 306
QY 120 GLDGMTCNLRKAKSFIHPILKKYIEAIOI--YGLATPLSVFDDLVSMNQTLAVLSLSP 178
Db 307 -----ENLVKTVGFVPKGLP-----SQPP 326
QY 179 SHVLEAQLNDSKEVLSVLYQMATTTEVLSSLGQKLLAFAGLSVLVLTGLFMKRFL-- 235
Db 327 PFHMDAVVNETTGEVLQ---EAQSFWDVSTLSSGLIIVP--LIALLETMAVVQAQADGK 381
QY 236 -----GPCWKYENIYITRFQVQDERERHOORPCVL-----PLNKEERRY 277
Db 382 PTDAQELTASGCN-----VANSFVQGLRSNGGIARGAILNASGVRTQLSNLYTSVI 434
QY 278 VIIPFWPTPKERKNLGLFPLPILHLCTIWLFAADVLL-YRLIFS--SKQFQSLPG 333
Db 435 VIALLYLTP-----CFYIIPAAIASI--IIAAVIFMWQYVIRKPMHSHKKTDLIPG 485

RESULT 15

PCT-US02-22217-17
; Sequence 17, Application PC/TUS0222217
; GENERAL INFORMATION:
; APPLICANT: MCGILL UNIVERSITY
; APPLICANT: DNA LANDMARKS, INC.
; TITLE OF INVENTION: NUCLEAR FERTILITY RESTORER GENES AND METHODS OF USE IN
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 16313-0135
; CURRENT APPLICATION NUMBER: PCT/US02/22217
; CURRENT FILING DATE: 2002-09-03
; PRIOR APPLICATION NUMBER: 60/305,026
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 60/305,363
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/308,736
; PRIOR FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 1494
; TYPE: PRT
; ORGANISM: Raphanus sativum
PCT-US02-22217-17

Query Match 4.2%; Score 102.5; DB 1; Length 1494;
Best Local Similarity 22.5%; Pred. No. 2.4;
Matches 78; Conservative 49; Mismatches 112; Indels 107; Gaps 18;
QY 103 TGVILGHVENIFHNFKGLLDGMTCNLRKAKSFIHPILKKYIEAIOIYGLATPLSVFD 162
Db 929 SGFISLKNVEDLVHFDGL-----CSINIRELNI---LROYHVDALSWI-----SRFD 973
QY 163 DLVSNQTLAVLSLSPSHVLEAQLNDSKEVLSVLYQMATTTEVLSSLGQKLLAFA-GLS 221

```

Db 974 DTM-----ADVREGKDO-RKLISDLSSLLQDASLGIOVSEIVEGLP 1014
Qy 222 LVLLGTGLFMKRFGLGPGCKMYENIYITRQVQFDERERHQORPCVLPNKEERRKYVII- 280
Db 1015 LV---EVELKK--ASCREKAQTVAARSSLDIFIQLISE-----AVWLQIEEEKLFVEIA 1064
Qy 281 -----PTWPTPKER-----KNLGLFPLPIHLHCIIWVLFPAVDYLLYRLIFSVSQOFQ 329
Db 1065 GTLSTARFW---EERASSILASETQMSDLKELVHFCVLVFL-----KGF- 1104
Qy 330 SLPGFEVHLKLHGEKQGTODIHDSSFNISVPEP-----NCIPKPKFLLSETWVPLSVI 383
Db 1105 -----VHOK-----DDVRMSVNICAVLPISLKGIENTIS-----LAETNLQNSEP 1144
Qy 384 LLIIVMLGLLSI-----LMQLKILVSASFYPSVERKRIQYLHAKLL 425
Db 1145 FLASVSSTASSPCSLLELPVLKDLVAQSKSLGVQLEEPRIETLTL 1190

```

Search completed: June 21, 2003, 13:34:17
Job time : 114 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2003, 13:23:47 ; Search time 308 Seconds
(without alignments)
983.846 Million cell updates/sec

Title: US-09-713-098-2

Perfect score: 2448

Sequence: 1 MGWTSGTDIFLSLWEIYVS.....LPVLKMKRQNDMASADKS 470

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA_Main:*

```

1: /cgn2_6/ptodata/1/paa/PCTUS COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06 COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07 COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08 COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09 COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US10 COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US11 COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US12 COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US13 COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US14 COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US15 COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US16 COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US17 COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US18 COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US19 COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US20 COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US21 COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US22 COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US23 COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US24 COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US25 COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US26 COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US27 COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US28 COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US29 COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US30 COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US31 COMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2448	100.0	470	1	PCT-US02-29964-459
2	2448	100.0	470	18	US-09-439-2552- <i>Abundant</i>
3	2448	100.0	470	21	US-09-713-098-2
4	2434	99.4	470	25	US-10-108-260A-3139
5	1799.5	73.5	352	1	PCT-US00-24008-98
6	1799.5	73.5	352	1	PCT-US01-11988-2122

```

7 1799.5 73.5 352 21 US-09-789-561-98
8 1799.5 73.5 352 22 US-09-833-245-2122
9 1314 53.7 291 1 PCT-US00-24008-142
10 1314 53.7 291 1 PCT-US01-11988-2119
11 1314 53.7 291 21 US-09-789-561-142
12 1314 53.7 291 22 US-09-833-245-2119
13 1314 53.7 291 23 US-09-950-083-3966
14 1314 53.7 291 25 US-10-105-299-4970
15 1314 53.7 292 1 PCT-US00-07535-85
16 1309 53.5 257 1 PCT-US00-24008-140
17 1309 53.5 257 1 PCT-US01-11988-2120
18 1309 53.5 257 1 PCT-US01-11988-2121
19 1309 53.5 257 21 US-09-789-561-140
20 1309 53.5 257 21 US-09-789-561-141
21 1309 53.5 257 22 US-09-833-245-2120
22 1309 53.5 257 22 US-09-833-245-2121
23 1309 53.5 257 23 US-09-950-083-3965
24 1309 53.5 257 25 US-10-105-299-4969
25 1309 53.5 258 1 PCT-US00-15136-80
26 1309 53.5 258 1 PCT-US01-16450A-2360
27 754.5 30.8 148 1 PCT-US01-16450A-2360
28 754.5 30.8 148 25 US-10-108-260A-4733
29 197 8.0 604 25 US-10-108-260A-4706
30 194 7.9 489 25 US-10-108-260A-4706
31 171 7.0 484 1 PCT-US02-08123-1065
32 171 7.0 484 1 PCT-US02-08124-448
33 171 7.0 484 1 PCT-US02-08278-1022
34 171 7.0 484 1 PCT-US02-08279-338
35 171 7.0 485 1 PCT-US99-05721-54
36 171 7.0 485 17 US-09-393-022-55
37 171 7.0 485 21 US-09-798-889-55
38 161.5 6.6 592 27 US-60-167-217-15054
39 161.5 6.6 592 27 US-60-173-464-13272
40 158.5 6.5 684 20 US-09-614-150-23619
41 158.5 6.5 684 27 US-60-173-464-19445
42 158.5 6.5 684 27 US-60-191-637-23702
43 158.5 6.5 684 27 US-60-191-681-18683
44 116 4.7 285 19 US-09-513-996A-71534
45 116 4.7 286 19 US-09-595-298A-525

```

ALIGNMENTS

```

RESULT 1
PCT-US02-29964-459
; Sequence 459, Application PC/TUS0229964
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wang, Dunrui
; APPLICANT: Ghosh, Malabika
; APPLICANT: Asundi, Vinod
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle W.
; APPLICANT: Weng, Gezhi
; APPLICANT: Haley-Vicente, Dana
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 809ACIP PCT
; CURRENT APPLICATION NUMBER: PCT/US02/29964
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/323,739
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/552,317

```

```

Sequence 98, Appl
Sequence 2122, Ap
Sequence 142, App
Sequence 2119, Ap
Sequence 142, App
Sequence 2119, Ap
Sequence 3966, Ap
Sequence 4970, Ap
Sequence 85, Appl
Sequence 140, App
Sequence 141, App
Sequence 2120, Ap
Sequence 2121, Ap
Sequence 140, App
Sequence 141, App
Sequence 2120, Ap
Sequence 2121, Ap
Sequence 3965, Ap
Sequence 4969, Ap
Sequence 80, Appl
Sequence 2360, Ap
Sequence 2360, Ap
Sequence 4733, Ap
Sequence 4706, Ap
Sequence 1065, Ap
Sequence 448, App
Sequence 1022, Ap
Sequence 338, App
Sequence 54, Appl
Sequence 55, Appl
Sequence 55, Appl
Sequence 15054, A
Sequence 12272, A
Sequence 23619, A
Sequence 19445, A
Sequence 23702, A
Sequence 18683, A
Sequence 71534, A
Sequence 525, App

```

; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: PCT/US00/35017
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: PCT/US01/03800
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 992
; SOFTWARE: dt_FL_genes Version 6.0
; SEQ ID NO 459
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-29964-459

Query Match 100.0%; Score 2448; DB 1; Length 470;
Best Local Similarity 100.0%; Pred. No. 3.5e-225;
Matches 470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGIWTSQTDIFLSLWEIYVSPRSGWMDFTQHLGVCCLVALISVGLLSVAACWFLPSIIA 60
DB 1 MGIWTSQTDIFLSLWEIYVSPRSGWMDFTQHLGVCCLVALISVGLLSVAACWFLPSIIA 60

QY 61 AAASWIIITCVLLCCSKHARCFILLVFLSCGLREGRNALIAAGTGVILGHVENIFHNFKG 120
DB 61 AAASWIIITCVLLCCSKHARCFILLVFLSCGLREGRNALIAAGTGVILGHVENIFHNFKG 120

QY 121 LLDGMCNLRKSFHSHPFLPKYIEAIOIYGLATPLSVFDDLVSNQTLAVSLFSPSH 180
DB 121 LLDGMCNLRKSFHSHPFLPKYIEAIOIYGLATPLSVFDDLVSNQTLAVSLFSPSH 180

QY 181 VLEAQNDSKGEVLSVLYQMATTTEVLSLQKLLAFAGLSVLGTGLFMKRFGLPCGW 240
DB 181 VLEAQNDSKGEVLSVLYQMATTTEVLSLQKLLAFAGLSVLGTGLFMKRFGLPCGW 240

QY 241 KYENIYITQFVQDERERHQPCVLPNKERRKYIIPFWPTPKERNKGLFLPI 300
DB 241 KYENIYITQFVQDERERHQPCVLPNKERRKYIIPFWPTPKERNKGLFLPI 300

QY 301 LIHLCIWLFAAVDYLRLIFSVSKQFSLPGFEVHLKLGKQGTQDIIHDSFSNISV 360
DB 301 LIHLCIWLFAAVDYLRLIFSVSKQFSLPGFEVHLKLGKQGTQDIIHDSFSNISV 360

QY 361 FEPCIPKPKFLSETWVPLSVILLVLMGLSSILMQILKILVSASFYPSVERKRIQYL 420
DB 361 FEPCIPKPKFLSETWVPLSVILLVLMGLSSILMQILKILVSASFYPSVERKRIQYL 420

QY 421 HAKLLKRSKQPLGEVKRRLSLYTKIHFWLPVLMKIRKQMDMASADKS 470
DB 421 HAKLLKRSKQPLGEVKRRLSLYTKIHFWLPVLMKIRKQMDMASADKS 470

RESULT 2
US-09-439-735-2
; Sequence 2, Application US/09439735
; GENERAL INFORMATION:
; APPLICANT: Zlot, Constance F.
; APPLICANT: Adema, Gosse J.
; APPLICANT: Figdor, Carl
; APPLICANT: Phillips, Joseph H.
; TITLE OF INVENTION: Mammalian Genes; Related Reagents and Methods
; FILE REFERENCE: DX1051
; CURRENT APPLICATION NUMBER: US/09/439,735
; CURRENT FILING DATE: 1999-11-15

; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 470
; TYPE: PRT
; ORGANISM: primate
US-09-439-735-2

Query Match 100.0%; Score 2448; DB 18; Length 470;
Best Local Similarity 100.0%; Pred. No. 3.5e-225;
Matches 470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGIWTSQTDIFLSLWEIYVSPRSGWMDFTQHLGVCCLVALISVGLLSVAACWFLPSIIA 60
DB 1 MGIWTSQTDIFLSLWEIYVSPRSGWMDFTQHLGVCCLVALISVGLLSVAACWFLPSIIA 60

QY 61 AAASWIIITCVLLCCSKHARCFILLVFLSCGLREGRNALIAAGTGVILGHVENIFHNFKG 120
DB 61 AAASWIIITCVLLCCSKHARCFILLVFLSCGLREGRNALIAAGTGVILGHVENIFHNFKG 120

QY 121 LLDGMCNLRKSFHSHPFLPKYIEAIOIYGLATPLSVFDDLVSNQTLAVSLFSPSH 180
DB 121 LLDGMCNLRKSFHSHPFLPKYIEAIOIYGLATPLSVFDDLVSNQTLAVSLFSPSH 180

QY 181 VLEAQNDSKGEVLSVLYQMATTTEVLSLQKLLAFAGLSVLGTGLFMKRFGLPCGW 240
DB 181 VLEAQNDSKGEVLSVLYQMATTTEVLSLQKLLAFAGLSVLGTGLFMKRFGLPCGW 240

QY 241 KYENIYITQFVQDERERHQPCVLPNKERRKYIIPFWPTPKERNKGLFLPI 300
DB 241 KYENIYITQFVQDERERHQPCVLPNKERRKYIIPFWPTPKERNKGLFLPI 300

QY 301 LIHLCIWLFAAVDYLRLIFSVSKQFSLPGFEVHLKLGKQGTQDIIHDSFSNISV 360
DB 301 LIHLCIWLFAAVDYLRLIFSVSKQFSLPGFEVHLKLGKQGTQDIIHDSFSNISV 360

QY 361 FEPCIPKPKFLSETWVPLSVILLVLMGLSSILMQILKILVSASFYPSVERKRIQYL 420
DB 361 FEPCIPKPKFLSETWVPLSVILLVLMGLSSILMQILKILVSASFYPSVERKRIQYL 420

QY 421 HAKLLKRSKQPLGEVKRRLSLYTKIHFWLPVLMKIRKQMDMASADKS 470
DB 421 HAKLLKRSKQPLGEVKRRLSLYTKIHFWLPVLMKIRKQMDMASADKS 470

RESULT 3
US-09-713-098-2
; Sequence 2, Application US/09713098
; GENERAL INFORMATION:
; APPLICANT: Zlot, Constance H.
; APPLICANT: Adema, Gosse J.
; APPLICANT: Figdor, Carl
; APPLICANT: Phillips, Joseph H.
; TITLE OF INVENTION: Mammalian Genes; Related Reagents and Methods
; FILE REFERENCE: DX1051Q
; CURRENT APPLICATION NUMBER: US/09/713,098
; CURRENT FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 470
; TYPE: PRT
; ORGANISM: primate
US-09-713-098-2

Query Match 100.0%; Score 2448; DB 21; Length 470;
Best Local Similarity 100.0%; Pred. No. 3.5e-225;
Matches 470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGIWTSQTDIFLSLWEIYVSPRSGWMDFTQHLGVCCLVALISVGLLSVAACWFLPSIIA 60
DB 1 MGIWTSQTDIFLSLWEIYVSPRSGWMDFTQHLGVCCLVALISVGLLSVAACWFLPSIIA 60

QY 61 AAASWITCVLLCCSKHARCFILLVFLSCGLREGNLIAGTGVILGHVENIHFNPKG 120
DB 61 AAASWITCVLLCCSKHARCFILLVFLSCGLREGNLIAGTGVILGHVENIHFNPKG 120
QY 121 LLDGTCNLRAKSFIIHPLKKYIEAIOIYGLATPLSVFDDVSNQTLAVSLFSPSH 180
DB 121 LLDGTCNLRAKSFIIHPLKKYIEAIOIYGLATPLSVFDDVSNQTLAVSLFSPSH 180
QY 181 VLEAQLNDSKGEVLSVLYOMATTTVEVLSLGGKLLAFAGLSLVLLGTGLFMKRFGLPCGW 240
DB 181 VLEAQLNDSKGEVLSVLYOMATTTVEVLSLGGKLLAFAGLSLVLLGTGLFMKRFGLPCGW 240
QY 241 KYENIYITRQVQFDERERHQRPCVPLNKEERKYYIIPFTWPTPKERNLGLFLLPI 300
DB 241 KYENIYITRQVQFDERERHQRPCVPLNKEERKYYIIPFTWPTPKERNLGLFLLPI 300
QY 301 LIHLICWVLFVAVDYLRLIFSFSKQFSLPGFVHLKLGKKGQGTQDIHDSSEFNISV 360
DB 301 LIHLICWVLFVAVDYLRLIFSFSKQFSLPGFVHLKLGKKGQGTQDIHDSSEFNISV 360
QY 361 FEPNCIPKPKFLLSETWVPLSVLLILVMLGLSSILMQLKILVSASFYSVERKRIQYL 420
DB 361 FEPNCIPKPKFLLSETWVPLSVLLILVMLGLSSILMQLKILVSASFYSVERKRIQYL 420
QY 421 HAKLLKRSKQPLGEVKRRRLSLYTKIHFWLPVLKMKIRKQMDMASADKS 470
DB 421 HAKLLKRSKQPLGEVKRRRLSLYTKIHFWLPVLKMKIRKQMDMASADKS 470

RESULT 4

US-10-108-260A-3139

; Sequence 3139, Application US/10108260A

; GENERAL INFORMATION:

; APPLICANT: HELIX RESEARCH INSTITUTE

; TITLE OF INVENTION: Novel full length cDNA

; FILE REFERENCE: H1-A0106

; CURRENT APPLICATION NUMBER: US/10/108,260A

; CURRENT FILING DATE: 2002-03-27

; NUMBER OF SEQ ID NOS: 5458

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3139

; LENGTH: 470

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-108-260A-3139

Query Match 99.4%; Score 2434; DB 25; Length 470;

Best Local Similarity 99.8%; Pred. No. 7.7e-224;

Matches 469; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGWITSGTDIFLSLWEIYVSPRSGWMDPIQHLGVCCVALLISVGLLSVAACWFLPSIIA 60
DB 1 MGWITSGTDIFLSLWEIYVSPRSGWMDPIQHLGVCCVALLISVGLLSVAACWFLPSIIA 60
QY 61 AAASWITCVLLCCSKHARCFILLVFLSCGLREGNLIAGTGVILGHVENIHFNPKG 120
DB 61 AAASWITCVLLCCSKHARCFILLVFLSCGLREGNLIAGTGVILGHVENIHFNPKG 120
QY 121 LLDGTCNLRAKSFIIHPLKKYIEAIOIYGLATPLSVFDDVSNQTLAVSLFSPSH 180
DB 121 LLDGTCNLRAKSFIIHPLKKYIEAIOIYGLATPLSVFDDVSNQTLAVSLFSPSH 180
QY 181 VLEAQLNDSKGEVLSVLYOMATTTVEVLSLGGKLLAFAGLSLVLLGTGLFMKRFGLPCGW 240
DB 181 VLEAQLNDSKGEVLSVLYOMATTTVEVLSLGGKLLAFAGLSLVLLGTGLFMKRFGLPCGW 240
QY 241 KYENIYITRQVQFDERERHQRPCVPLNKEERKYYIIPFTWPTPKERNLGLFLLPI 300
DB 241 KYENIYITRQVQFDERERHQRPCVPLNKEERKYYIIPFTWPTPKERNLGLFLLPI 300
QY 301 LIHLICWVLFVAVDYLRLIFSFSKQFSLPGFVHLKLGKKGQGTQDIHDSSEFNISV 360
DB 301 LIHLICWVLFVAVDYLRLIFSFSKQFSLPGFVHLKLGKKGQGTQDIHDSSEFNISV 360

QY 361 FEPNCIPKPKFLLSETWVPLSVLLILVMLGLSSILMQLKILVSASFYSVERKRIQYL 420
DB 361 FEPNCIPKPKFLLSETWVPLSVLLILVMLGLSSILMQLKILVSASFYSVERKRIQYL 420
QY 421 HAKLLKRSKQPLGEVKRRRLSLYTKIHFWLPVLKMKIRKQMDMASADKS 470
DB 421 HAKLLKRSKQPLGEVKRRRLSLYTKIHFWLPVLKMKIRKQMDMASADKS 470

RESULT 5

PCT-US00-24008-98

; Sequence 98, Application PC/TUS0024008

; GENERAL INFORMATION:

; APPLICANT: Human Genome Sciences, Inc.

; TITLE OF INVENTION: 52 Human secreted proteins

; FILE REFERENCE: P2043PCT

; CURRENT APPLICATION NUMBER: PCT/US00/24008

; CURRENT FILING DATE: 2000-08-31

; PRIOR APPLICATION NUMBER: 60/152,317

; PRIOR FILING DATE: 1999-09-03

; PRIOR APPLICATION NUMBER: 60/152,315

; PRIOR FILING DATE: 1999-09-03

; NUMBER OF SEQ ID NOS: 194

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 98

; LENGTH: 352

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (284)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

PCT-US00-24008-98

Query Match 73.5%; Score 1799.5; DB 1; Length 352;

Best Local Similarity 96.4%; Pred. No. 3.7e-163;

Matches 344; Conservative 2; Mismatches 4; Indels 7; Gaps 1;

QY 27 MDFOIHLGVCCVALLISVGLLSVAACWFLPSIIATAAASWITCVLLCCSKHARCFILLVF 86
DB 1 MDFOIHLGVCCVALLISVGLLSVAACWFLPSIIATAAASWITCVLLCCSKHARCFILLVF 60
QY 87 LSCGLREGNLIAGTGVILGHVENIHFNFKGLDGMTCNLRKSFHIFPLKKYIE 146
DB 61 LSCGLREGNLIAGTGVILGHVENIHFNFKGLDGMTCNLRKSFHIFPLKKYIE 120
QY 147 AIOIYGLATPLSVFDDVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYOMATTEV 206
DB 121 AIOIYGLATPLSVFDDVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYOMATTEV 180
QY 207 LSSLGOKLLAFAGLSLVLLGTGLFMKRFGLPCGWKYNIIYITRQVQFDERERHQRPCV 266
DB 181 LSSLGOKLLAFAGLSLVLLGTGLFMKRFGLPCGWKYNIIYITRQVQFDERERHQRPCV 240
QY 267 LPNKEERKYYIIPFTWPTPKERNLGLFLLIHLICWVLFVAVDYLRLIFSFSK 326
DB 241 LPNKEERKYYIIPFTWPTPKERNLGLFLLIHLICWVLFVAVDYLRLIFSFSK 300
QY 327 QFQSLPGFVHLKLGKKGQGTQDIHDSSEFNISVFEPCNIPKPKFLLSETWVPLSVI 383
DB 301 QFQSLPGFVHLKLGKKGQGTQDIHDSSEFNISVFEPCNIPKPKFLLSETWVPLSVI 350

RESULT 6

PCT-US01-11988-2122

; Sequence 2122, Application PC/TUS0111988

; GENERAL INFORMATION:

; APPLICANT: Human Genome Sciences, Inc.

; TITLE OF INVENTION: Albumin Fusion Proteins

; FILE REFERENCE: PF546PCT

; CURRENT APPLICATION NUMBER: PCT/US01/11988

; CURRENT FILING DATE: 2001-01-12

Query Match	73.5%;	Score 1799.5;	DB 21;	Length 352;
Best Local Similarity	96.4%;	Pred. No. 3.7e-163;		
Matches 344;	Conservative 2;	Mismatches 4;	Indels 7;	Gaps 1;
QY	27	MDFIQHLGVCCCLVALISVGLLSVAACWFLPSIIIAAASWIIITCVLLCCSKHARC	FILLVF	86
DB	1	MDFIQHLGVCCCLVALISVGLLSVAACWFLPSIIIAAASWIIITCVLLCCSKHARC	FILLVF	60
QY	87	LSCGLREGRNALIAAGTGIVILGHVENIIFHNFKGLLDGMTCNLRAKSFSIHFP	LLKKYIE	146
DB	61	LSCGLREGRNALIAAGTGIVILGHVENIIFHNFKGLLDGMTCNLRAKSFSIHFP	LLKKYIE	120
QY	147	AIQWYIGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQ	MATTEV	206
DB	121	AIQWYIGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQ	MATTEV	180
QY	207	LSSLGQKLAFAGLSVLIGTGLFMKRFGLPCGCKYENIITROFVQFDERERH	QOORPCV	266
DB	181	LSSLGQKLAFAGLSVLVLTGLFMKRFGLPCGCKYENIITROFVQFDERERH	QOORPCV	240
QY	267	LPNKEERKYYIIPTFWPTPKERKNLGLFFLPIHLHCIIWVLF	FAADVLYLLYRLFVSVK	326
DB	241	LPNKEERKYYIIPTFWPTPKERKNLGLFFLPIHLHCIIWVLF	FAADVLYLLYRLFVSVK	300
QY	327	QFQSLPGFGEVHLKHGEKQGTQDIHDSFNISVFPEPCIPKPKFLLSETW	VPLSVI	383
DB	301	QFQSLPGFGEVHLKHGEKQGTQDIHDSFNISVFPEPCIPK	-----WQALKLL	350
RESULT 8				
US-09-833-245-2122				
; Sequence 2122, Application US/09833245				
; GENERAL INFORMATION:				
; APPLICANT: Human Genome Sciences, Inc.				
; TITLE OF INVENTION: PF546PCT				
; FILE REFERENCE: PF546PCT				
; CURRENT APPLICATION NUMBER: US/09/833,245				
; CURRENT FILING DATE: 2001-04-12				
; PRIOR APPLICATION NUMBER: 60/229, 358				
; PRIOR FILING DATE: 2000-04-12				
; PRIOR APPLICATION NUMBER: 60/256, 931				
; PRIOR FILING DATE: 2000-12-21				
; PRIOR APPLICATION NUMBER: 60/199, 384				
; PRIOR FILING DATE: 2000-04-25				
; NUMBER OF SEQ ID NOS: 2267				
; SOFTWARE: PatentIn Ver. 2.1				
; SEQ ID NO 2122				
; LENGTH: 352				
; TYPE: PRT				
; ORGANISM: Homo sapiens				
; FEATURE:				
; NAME/KEY: SITE				
; LOCATION: (284)				
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids				
US-09-833-245-2122				
Query Match	73.5%;	Score 1799.5;	DB 22;	Length 352;
Best Local Similarity	96.4%;	Pred. No. 3.7e-163;		
Matches 344;	Conservative 2;	Mismatches 4;	Indels 7;	Gaps 1;
QY	27	MDFIQHLGVCCCLVALISVGLLSVAACWFLPSIIIAAASWIIITCVLLCCSKHARC	FILLVF	86
DB	1	MDFIQHLGVCCCLVALISVGLLSVAACWFLPSIIIAAASWIIITCVLLCCSKHARC	FILLVF	60
QY	87	LSCGLREGRNALIAAGTGIVILGHVENIIFHNFKGLLDGMTCNLRAKSFSIHFP	LLKKYIE	146
DB	61	LSCGLREGRNALIAAGTGIVILGHVENIIFHNFKGLLDGMTCNLRAKSFSIHFP	LLKKYIE	120
QY	147	AIQWYIGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQ	MATTEV	206
DB	121	AIQWYIGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQ	MATTEV	180

QY 207 LSSLGOKLLAFAGLSLVLLGTGLFMKRFGLGPCGKWKYENIYITRQFVQFDERERHQORPCV 266
Db |||||
181 LSSLGOKLLAFAGLSLVLLGTGLFMKRFGLGPCGKWKYENIYITRQFVQFDERERHQORPCV 240
QY 267 LPLNKEERKYYIIPTFWPTPKERKNLGLFFLPILHLICI 306
Db |||||
241 LPLNKEERK-----NKELKILSM-ILP-LIYLCL 268

RESULT 12

US-09-833-245-2119
; Sequence 2119, Application US/09833245
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PFS46PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2119
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-245-2119

Query Match 53.7%; Score 1314; DB 22; Length 291;

Best Local Similarity 92.5%; Pred. No. 1e-116;
Matches 259; Conservative 4; Mismatches 5; Indels 12; Gaps 3;

QY 27 MDFIQLGVCCLVALLISVLLSVAACWFLPSIIAAAASWITCVLLCCSKHARCFILLVF 86
Db 1 MDFIQLGVCCLVALLISVLLSVAACWFLPSIIAAAASWITCVLLCCSKHARCFILLVF 60
QY 87 LSCGLREGNALIAAGTGIVILGHVENIHFNPKGLLDGMTCNLRKSFHIFPLKKYIE 146
Db |||||
61 LSCGLREGNALIAAGTGIVILGHVENIHFNPKGLLDGMTCNLRKSFHIFPLKKYIE 120
QY 147 ALQWYIGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLQMATTTVEV 206
Db |||||
121 ALQWYIGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLQMATTTVEV 180
QY 207 LSSLGOKLLAFAGLSLVLLGTGLFMKRFGLGPCGKWKYENIYITRQFVQFDERERHQORPCV 266
Db |||||
181 LSSLGOKLLAFAGLSLVLLGTGLFMKRFGLGPCGKWKYENIYITRQFVQFDERERHQORPCV 240
QY 267 LPLNKEERKYYIIPTFWPTPKERKNLGLFFLPILHLICI 306
Db |||||
241 LPLNKEERK-----NKELKILSM-ILP-LIYLCL 268

RESULT 13

US-09-950-083-3966
; Sequence 3966, Application US/09950083
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS805
; CURRENT APPLICATION NUMBER: US/09/950,083
; CURRENT FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US01/11988
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US00/06043

; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06012
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06058
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06044
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06059
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06042
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06014
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06013
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06049
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06057
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06824
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: PCT/US00/06765
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: PCT/US00/06792
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: PCT/US00/06830
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: PCT/US00/06782
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: PCT/US00/06822
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: PCT/US00/06791
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: PCT/US00/06828
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: PCT/US00/06823
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: PCT/US00/06781
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: PCT/US00/07505
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/US00/07440
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/US00/07506
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/US00/07507
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/US00/07535
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/US00/07525
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/US00/07534
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/US00/07483
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/US00/07526
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/US00/07527
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/US00/07661
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/US00/07579
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/US00/07723
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/US00/07724
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/US00/14929
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/07722
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/US00/07578
; PRIOR FILING DATE: 2000-03-23

; PRIOR APPLICATION NUMBER: PCT/US00/07726
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/US00/07677
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/US00/07725
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/US00/09070
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/US00/08982
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/US00/08983
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/US00/09067
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/US00/09066
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/US00/09068
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/US00/08981
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/US00/08980
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/US00/09071
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/US00/09069
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/US00/15136
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/14926
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/14963
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/15135
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/14934
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/14933
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/15137
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/14928
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/14973
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/14964
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/26376
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: PCT/US00/26371
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: PCT/US00/26324
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: PCT/US00/26323
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: PCT/US00/26337
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: PCT/US01/13318
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/124,146
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/167,061
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/124,093
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/166,989
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/124,145
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/168,654
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/124,099
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/168,661

; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/124,096
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/168,622
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/124,143
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/168,663
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/124,095
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/138,598
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: US 60/168,665
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/125,360
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: US 60/138,626
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: US 60/168,662
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/124,144
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/138,574
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: US 60/168,667
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/124,142
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/138,597
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: US 60/168,666
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/125,359
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: US 60/168,664
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/126,051
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/169,906

Query Match 53.7%; Score 1314; DB 23; Length 291;
Best Local Similarity 92.5%; Pred. No. 1e-116;
Matches 259; Conservative 4; Mismatches 5; Indels 12; Gaps 3;

Qy	27	MDFIQHLGVCCLVALLISVGLLSVAACWFLPSIIAAAAAIIITCVLLCCSKHARCFILLVF	86
Db	1	MDFIQHLGVCCLVALLISVGLLSVAACWFLPSIIAAAAAIIITCVLLCCSKHARCFILLVF	60
Qy	87	LSCGLREGRNALIAAGTGIVILGHVENI FHNFKGLLDGMTCNLRKSFSIHFPLKKYIE	146
Db	61	LSCGLREGRNALIAAGTGIVILGHVENI FHNFKGLLDGMTCNLRKSFSIHFPLKKYIE	120
Qy	147	AIQWIYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTVE	206
Db	121	AIQWIYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTVE	180
Qy	207	LSSLGOKLLAFAGLSLVLLGTGLFMKRFGLGPCWKYENIYITRQVQVDFRERHQORPCV	266
Db	181	LSSLGOKLLAFAGLSLVLLGTGLFMKRFGLGPCWKYENIYITRQVQVDFRERHQORPCV	240
Qy	267	LPLNKEERKVIITPFWPTPKERNLGLFFLPILHLCI	306
Db	241	LPLNKEERK-----NKELKLSM-IIP-LIYLCL	268

RESULT 14

US-10-105-298-4970
; Sequence 4970, Application US/10105299
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS950

```
; CURRENT APPLICATION NUMBER: US/10/105,299
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 15197
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4970
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-105-299-4970

Query Match          53.7%; Score 1314; DB 25; Length 291;
Best Local Similarity 92.5%; Pred. No. 1.le-116;
Matches 259; Conservative 4; Mismatches 5; Indels 12; Gaps 3;

QY 27 MDFIOHGLGVCCLVALISVGLLSVAACWFLPSIIAAAASWIIITCVLLCCSKHARCFILLVF 86
DB 1 MDFIOHGLGVCCLVALISVGLLSVAACWFLPSIIAAAASWIIITCVLLCCSKHARCFILLVF 60

QY 87 LSCGLREGNALLIAAGTGIVILGHVENIFHNFKGLDGMTCNLRKSFSIHFPPLKKYIE 146
DB 61 LSCGLREGNALLIAAGTGIVILGHVENIFHNFKGLDGMTCNLRKSFSIHFPPLKKYIE 120

QY 147 AIQWYIGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 206
DB 121 AIQWYIGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 180

QY 207 LSSLGOKLLAFAGLSLVLLGTGLFMKRFLGPCGKWKYENIYITRQVFQFDERERHQRPCV 266
DB 181 LSSLGOKLLAFAGLSLVLLGTGLFMKRFLGPCGKWKYENIYITRQVFQFDERERHQRPCV 240

QY 267 LPLNKEERRKYVIPTFWPTPKERNKLGFLFPIHLICI 306
DB 241 LPLNKEERRK-----NKEKILSM-ILP-LIYLCL 268
```

Search completed: June 21, 2003, 13:32:18
Job time : 310 secs

```
RESULT 15
PCT-US00-07535-85
; Sequence 85, Application PC/TUS0007535
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 50 Human Secreted Proteins
; FILE REFERENCE: PS525PCT
; CURRENT APPLICATION NUMBER: PCT/US00/07535
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/126,511
; EARLIER FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: 60/172,413
; EARLIER FILING DATE: 1999-12-17
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (292)
; OTHER INFORMATION: Xaa equals stop translation
PCT-US00-07535-85
```

```
Query Match          53.7%; Score 1314; DB 1; Length 292;
Best Local Similarity 92.5%; Pred. No. 1.le-116;
Matches 259; Conservative 4; Mismatches 5; Indels 12; Gaps 3;

QY 27 MDFIOHGLGVCCLVALISVGLLSVAACWFLPSIIAAAASWIIITCVLLCCSKHARCFILLVF 86
DB 1 MDFIOHGLGVCCLVALISVGLLSVAACWFLPSIIAAAASWIIITCVLLCCSKHARCFILLVF 60

QY 87 LSCGLREGNALLIAAGTGIVILGHVENIFHNFKGLDGMTCNLRKSFSIHFPPLKKYIE 146
DB 61 LSCGLREGNALLIAAGTGIVILGHVENIFHNFKGLDGMTCNLRKSFSIHFPPLKKYIE 120
```

```
QY 147 AIQWYIGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 206
DB 121 AIQWYIGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 180

QY 207 LSSLGOKLLAFAGLSLVLLGTGLFMKRFLGPCGKWKYENIYITRQVFQFDERERHQRPCV 266
DB 181 LSSLGOKLLAFAGLSLVLLGTGLFMKRFLGPCGKWKYENIYITRQVFQFDERERHQRPCV 240

QY 267 LPLNKEERRKYVIPTFWPTPKERNKLGFLFPIHLICI 306
DB 241 LPLNKEERRK-----NKEKILSM-ILP-LIYLCL 268
```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2003, 13:21:27 ; Search time 42 Seconds
(without alignments)
1075.790 Million cell updates/sec

Title: US-09-713-098-2

Perfect score: 2448

Sequence: 1 MGWTSGTDIFLSLWEIYVS.....LPVLKMRKKQMDASADKS 470

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	113.5	4.6	351	T15112	hypothetical prote
2	102.5	4.2	407	E71665	bicyclomycin resis
3	102	4.2	391	H81265	probable transmemb
4	101	4.1	645	T28867	hypothetical prote
5	100.5	4.1	729	I52481	PEPT 2 - human
6	100	4.1	997	A40812	Ca2+-transporting
7	100	4.1	997	S04651	Ca2+-transporting
8	100	4.1	1042	B40812	Ca2+-transporting
9	100	4.1	1042	S04652	Ca2+-transporting
10	99.5	4.1	470	G8516	dicarboxylase tran
11	99.5	4.1	690	G84638	hypothetical prote
12	99	4.0	510	A30594	Ca2+-transporting
13	99	4.0	997	PWRBSC	Ca2+-transporting
14	99	4.0	997	B31981	Ca2+-transporting
15	99	4.0	997	B31982	Ca2+-transporting
16	99	4.0	997	S23444	Ca2+-transporting
17	99	4.0	1042	PWRBMC	Ca2+-transporting
18	99	4.0	1042	A31981	Ca2+-transporting
19	99	4.0	1042	A33881	Ca2+-transporting
20	99	4.0	1043	A31982	Ca2+-transporting
21	98	4.0	573	AG2247	hypothetical prote
22	97.5	4.0	1065	E69795	acriflavin resista
23	97	4.0	592	H97105	carbon starvation
24	97	4.0	614	A69845	Na+/H+ antiporter
25	97	4.0	626	D70178	PTS system, fructo
26	96.5	3.9	732	T23505	hypothetical prote
27	96	3.9	405	D64522	carboxynorspermi
28	96	3.9	451	G71561	probable integral
29	96	3.9	514	S56384	hypothetical 56.3K

transmembrane effl
NADH2 dehydrogenas
probable ABC trans
probable transport
probable transport
disease resistance
tellurite resistanc
NADH dehydrogenase
probable V-type AT
hypothetical prote
alkaline phosphata
hypothetical prote
neurotransmitter r
conserved hypothet
hypothetical prote
probable NADH dehy

ALIGNMENTS

RESULT 1

T15112

hypothetical protein ZC132.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jan-2000

C:Accession: T15112

R:Bradshaw, H.; Devlin, K.

submitted to the EMBL Data Library, July 1997

A:Description: The sequence of C. elegans cosmid ZC132.

A:Reference number: Z18294

A:Accession: T15112

A:Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: DNA

A:Residues: 1-351 <BRA>

A:Cross-references: EMBL:AF014939; NID:g2275620; PID:g2275626; PIDN:AAB63929.1; GSPDB:G

A:Experimental source: strain Bristol N2; clone ZC132

C:Genetics:

A:Gene: CBSP:ZC132.7

A:Map position: 5

A:Introns: 118/1; 181/3

C:Superfamily: Caenorhabditis hypothetical protein C49G7.2

Query Match	4.6%;	Score 113.5;	DB 2;	Length 351;
Best Local Similarity	20.2%;	Pred. No. 0.059;		
Matches	73;	Conservative 57;	Mismatches 111;	Indels 121; Gaps 15;
QY	133	SFSIHF-----	PLLKYYIEAQW-IYGLATPLSVDPDDLVSVMNQTLAVSLFSPSHV	181
DB	37	SIPIHFGMYCIYKTPVMK-----	TVKMYLFALHVMIIAFD-----	YFSFSLTAPFL 85
QY	182	LEAQLNDSKEVLSVLVYQMATTEVLSLQKLLAFAGLSVLVLTGTLFMKRFILGPCGWK	241	
DB	86	LIPKLG--GYILGILKYTSMPLDYLTISVMGAYMGISIV----	SIFENRFYIVCDFA	138
QY	242	YENIYITROFQVDERERHQORPCVLPNKEERKVVIIPTF-----	WPTPKERKNLGLF	296
DB	139	FKNHVVLRRIWL-----	ATHVIVTFLTPVFLIPDQKIAPLM	179
QY	297	F-----LPILI-----	HLCIWL-----	FAVDYLLYRLIFSRSKQF 328
DB	180	FQKLPCLPSIYEAPILVLSSELYTHATISVVYIFLVLFVIESFIFVGLFINIVQMKEHK	239	
QY	329	QSLPGFEVHLKHXGKQGTQDIHDSFSNISVPEPNCIPKPKLLSTWPLSVLLILV	388	
DB	240	MSPKTPELOKK-----	FIIT-----	LLIQVSIPIICFIETLI 271
QY	389	MLGLSSILMQLKILVSA-----	SFYPSVERKEIOYLHA-----	KLLKRSKQPLGEV 436
DB	272	YIGFAVLINYNQGLNNAITLAFSCHGSTIALIALHAPYREVAQDLRLKLRMSPVVEV	331	
QY	437	KR	438	

Db 332 SQ 333

RESULT 2

E71665

bicyclomycin resistance protein (bcr1) RP603 - Rickettsia prowazekii

C:Species: Rickettsia prowazekii

C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000

C:Accession: E71665

R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark, U.

Nature 396, 133-140, 1998

A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.

A:Reference number: A71630; MUID:99039499; PMID:9823893

A:Accession: E71665

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-407 <AND>

A:Cross-references: GB:AJ235272; GB:AJ235269; NID:G3861033; PIDN:CAA15047.1; PID:G386114

A:Experimental source: strain Madrid E

C:Genetics:

A:Gene: bcr1; RP603

C:Superfamily: bicyclomycin resistance protein

Query Match 4.2%; Score 102.5; DB 2; Length 407;

Best Local Similarity 16.3%; Pred. No. 0.62;

Matches 79; Conservative 79; Mismatches 131; Indels 195; Gaps 22;

QY 58 IIAAAAWIITCVLLC-----CSKHARCFIL-----LVFL--SCGLRE- 93

Db 3 IIAKIPANMLLCFILSPTEIYISGLSLTRCFSDICITQMTSLTFLGFAVGILSL 62

QY 94 -----GNALIAAGTGVILGHVENIFHNFKGLLDGTMCLNRAKSFISHPFLKKYIE 146

Db 63 GRLSDIYGRRPVLLGLFIYVSSISIF-----SFNIEMLMIARFIQ 105

QY 147 A-----IQWYGLATPLSVF-----DDLVSNQTLA 172

Db 106 AFGSVSGVIGSMARDSYQGAELSYVYAILSPWLLFIPALGSYIGGYIIEVLSWHYVFI 165

QY 173 VSLFSPSHVLEAQLNDSKGEVLSVLYQMATTEVLSLQKLLAPAGLSLGLTGLPMK 232

Db 166 ---FFSLAGTI-----LIALYQILPETYIYAFQSSKYFEVFNIIKDKMLWLY 213

QY 233 RF-LGPGCKWYENIYITROFODERERHQRPCVLPNKEERRKVKVIITPFWPT----- 286

Db 214 ARIIGAFNGIYGFIEAPFILLDQNR-----VLPSFYGLAFLL 253

QY 287 -----PKERNLGLFPLIHLICIVWLFPA-----VDLYLRLL 321

Db 254 SPASIFGGFLGYLIKQKQVYDKVMISGIF-----SLCGCILFAVDGFILFVLSNV 308

QY 322 FSVSKQFQSLPGFEVH-----LKLHGEKQGTQDIHDSFNISVFPFN-C 365

Db 309 FAIAMIFFPM-----MTHIGHSLIIATLTLYALEDYATVTGTAGSIFGAIYVVIASVTVC 365

QY 366 IKPKRFLSETWPLSVILLIIVMLGLLSILMQLKILVSASFYPSVERKRIQYLHAKLL 425

Db 366 VSK---IHGETISNLSLLCLVL-----SISSVI-----SPY-----YI-CLHY 399

QY 426 KKRS 429

Db 400 KKKS 403

RESULT 3

H81265

probable transmembrane transport protein Cj1684c [imported] - Campylobacter jejuni (stra

C:Species: Campylobacter jejuni

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002

C:Accession: H81265

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling

C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell

Nature 403, 665-668, 2000

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: H81265

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-391 <PAR>

A:Cross-references: GB:AL111168; GB:AL139079; GB:AL111168; NID:96968971; PIDN:CA873670.1; PID:969690

A:Experimental source: serotype O2, strain NCTC 11168

C:Genetics:

A:Gene: Cj1684c

Query Match 4.1%; Score 101; DB 2; Length 645;

Best Local Similarity 24.1%; Pred. No. 1.5;

Matches 65; Conservative 46; Mismatches 95; Indels 64; Gaps 16;

C:Superfamily: Caenorhabditis elegans hypothetical protein T26H2.7

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: H81265

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-645 <BRA>

A:Cross-references: EMBL:U50300; PIDN:AAC48106.1; GSPDB:GN00023; CESP:R03H4.5

A:Experimental source: strain Bristol N2; clone R03H4

C:Genetics:

A:Gene: CESP:R03H4.5

A:Map position: 5

A:Introns: 35/2; 66/2; 123/3; 166/3; 279/3; 320/1; 422/3; 463/1; 499/2; 527/2; 590/3

C:Superfamily: Caenorhabditis elegans hypothetical protein T26H2.7

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: H81265

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-645 <BRA>

A:Cross-references: EMBL:U50300; PIDN:AAC48106.1; GSPDB:GN00023; CESP:R03H4.5

A:Experimental source: strain Bristol N2; clone R03H4

C:Genetics:

A:Gene: CESP:R03H4.5

A:Map position: 5

A:Introns: 35/2; 66/2; 123/3; 166/3; 279/3; 320/1; 422/3; 463/1; 499/2; 527/2; 590/3

C:Superfamily: Caenorhabditis elegans hypothetical protein T26H2.7

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: H81265

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-645 <BRA>

A:Cross-references: EMBL:U50300; PIDN:AAC48106.1; GSPDB:GN00023; CESP:R03H4.5

A:Experimental source: strain Bristol N2; clone R03H4

C:Genetics:

A:Gene: CESP:R03H4.5

A:Map position: 5

A:Introns: 35/2; 66/2; 123/3; 166/3; 279/3; 320/1; 422/3; 463/1; 499/2; 527/2; 590/3

C:Superfamily: Caenorhabditis elegans hypothetical protein T26H2.7

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: H81265

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-645 <BRA>

A:Cross-references: EMBL:U50300; PIDN:AAC48106.1; GSPDB:GN00023; CESP:R03H4.5

A:Experimental source: strain Bristol N2; clone R03H4

C:Genetics:

A:Gene: CESP:R03H4.5

A:Map position: 5

A:Introns: 35/2; 66/2; 123/3; 166/3; 279/3; 320/1; 422/3; 463/1; 499/2; 527/2; 590/3

C:Superfamily: Caenorhabditis elegans hypothetical protein T26H2.7

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: H81265

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-645 <BRA>

A:Cross-references: EMBL:U50300; PIDN:AAC48106.1; GSPDB:GN00023; CESP:R03H4.5

A:Experimental source: strain Bristol N2; clone R03H4

C:Genetics:

A:Gene: CESP:R03H4.5

A:Map position: 5

A:Introns: 35/2; 66/2; 123/3; 166/3; 279/3; 320/1; 422/3; 463/1; 499/2; 527/2; 590/3

C:Superfamily: Caenorhabditis elegans hypothetical protein T26H2.7

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: H81265

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-645 <BRA>

A:Cross-references: EMBL:U50300; PIDN:AAC48106.1; GSPDB:GN00023; CESP:R03H4.5

A:Experimental source: strain Bristol N2; clone R03H4

C:Genetics:

A:Gene: CESP:R03H4.5

A:Map position: 5

A:Introns: 35/2; 66/2; 123/3; 166/3; 279/3; 320/1; 422/3; 463/1; 499/2; 527/2; 590/3

C:Superfamily: Caenorhabditis elegans hypothetical protein T26H2.7

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: H81265

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-645 <BRA>

A:Cross-references: EMBL:U50300; PIDN:AAC48106.1; GSPDB:GN00023; CESP:R03H4.5

A:Experimental source: strain Bristol N2; clone R03H4

C:Genetics:

A:Gene: CESP:R03H4.5

A:Map position: 5

A:Introns: 35/2; 66/2; 123/3; 166/3; 279/3; 320/1; 422/3; 463/1; 499/2; 527/2; 590/3

C:Superfamily: Caenorhabditis elegans hypothetical protein T26H2.7

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: H81265

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-645 <BRA>

A:Cross-references: EMBL:U50300; PIDN:AAC48106.1; GSPDB:GN00023; CESP:R03H4.5

A:Experimental source: strain Bristol N2; clone R03H4

C:Genetics:

A:Gene: CESP:R03H4.5

A:Map position: 5

A:Introns: 35/2; 66/2; 123/3; 166/3; 279/3; 320/1; 422/3; 463/1; 499/2; 527/2; 590/3

C:Superfamily: Caenorhabditis elegans hypothetical protein T26H2.7

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: H81265

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-645 <BRA>

A:Cross-references: EMBL:U50300; PIDN:AAC48106.1; GSPDB:GN00023; CESP:R03H4.5

A:Experimental source: strain Bristol N2; clone R03H4

C:Genetics:

A:Gene: CESP:R03H4.5

A:Map position: 5

A:Introns: 35/2; 66/2; 123/3; 166/3; 279/3; 320/1; 422/3; 463/1; 499/2; 527/2; 590/3

C:Superfamily: Caenorhabditis elegans hypothetical protein T26H2.7

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: H81265

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-645 <BRA>

A:Cross-references: EMBL:U50300; PIDN:AAC48106.1; GSPDB:GN00023; CESP:R03H4.5

A:Experimental source: strain Bristol N2; clone R03H4

C:Genetics:

A:Gene: CESP:R03H4.5

A:Map position: 5

A:Introns: 35/2; 66/2; 123/3; 166/3; 279/3; 320/1; 422/3; 463/1; 499/2; 527/2; 590/3

C:Superfamily: Caenorhabditis elegans hypothetical protein T26H2.7

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: H81265

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-645 <BRA>

A:Cross-references: EMBL:U50300; PIDN:AAC48106.1; GSPDB:GN00023; CESP:R03H4.5

A:Experimental source: strain Bristol N2; clone R03H4

C:Genetics:

A:Gene: CESP:R03H4.5

A:Map position: 5

A:Introns: 35/2; 66/2; 123/3; 166/3; 279/3; 320/1; 422/3; 463/1; 499/2; 527/2; 590/3

C:Superfamily: Caenorhabditis elegans hypothetical protein T26H2.7

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: H81265

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-645 <BRA>

A:Cross-references: EMBL:U50300; PIDN:AAC48106.1; GSPDB:GN00023; CESP:R03H4.5

A:Experimental source: strain Bristol N2; clone R03H4

C:Genetics:

A:Gene: CESP:R03H4.5

A:Map position: 5

A:Introns: 35/2; 66/2; 123/3; 166/3; 279/3; 320/1; 422/3; 463/1; 499/2; 527/2; 590/3

C:Superfamily: Caenorhabditis elegans hypothetical protein T26H2.7

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: H81265

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-645 <BRA>

A:Cross-references: EMBL:U50300; PIDN:AAC48106.1; GSPDB:GN00023; CESP:R03H4.5

A:Experimental source: strain Bristol N2; clone R03H4

C:Genetics:

A:Gene: CESP:R03H4.5

A:Map position: 5

A:Introns: 35/2; 66/2; 123/3; 166/3; 279/3; 320/1; 422/3; 463/1; 499/2; 527/2; 590/3

C:Superfamily: Caenorhabditis elegans hypothetical protein T26H2.7

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy

A:Reference

```

Qy 215 LAFAGLSVLVLGTLGFMKRFGLGPCGWKYNFY-----ITRFVQFDERERHQQ-RPCV 266
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 73 LYFLAIFLILILVHVFLPDFL---WQNNRYSFASFLITNQLVTHDQSYDFNEFVSV 128
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 267 LPLN-----KSERRKYVLIPTFWPTPKERKNLGLFLP-----LLIHLCIWVLFAAVD 314
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 129 SSMNAFLHLWSLSEVMQFYLLVPFIF-----LGIQFLKNDTLKLAVSLMTVFGFTG 180
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 315 YLL-----YRLIFSVSKQFOSLPQFEV--HLKLGKRGQGTQDIIDSSFNISVPEPNCI 366
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 181 FALILDKFAFNFLFLRWQFSS--GFIVLFWTKIRSRKPPNK---NDSEKDISTIS---I 232
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 367 PKPKPLSETWPIVSVILLIIVMLGLSSILMQLKILVSAGYPVSERRKRIQYLHAKLLK 426
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 233 PIKDFW---VALSIILGLSLPKBEINVLRLPLVTLATA-FIIGAESKNVQLLNSKVLV 287
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 427 KRSQPLGEVKERKRLSLYLTKIHFVLPVLKM 456
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 288 Y-----IGDIS-----YVVYLWHM-PPVISI 306
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 5
I52481
PEPT 2 - human
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C:Accession: I52481
E:Liu, W.; Liang, R.; Ramamoorthy, S.; Fei, Y.J.; Ganapathy, M.E.; Hediger, M.I.
Biochim. Biophys. Acta 1235, 461-466, 1995
A:Title: Molecular cloning of PEPT 2, a new member of the H+/peptide cotranspor-
A:Reference number: I52481; MUID:95275926; PMID:7756356
A:Accession: I52481
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-729 <RES>
A:Cross-references: GB:S78203; NID:g999212; PIDN:AAB34388.1; PID:g999213
C:Superfamily: peptide transport protein PEPT1

```

Query Match	4.1%;	Score 100.5;	DB 2;	Length 729;
Best Local Similarity	19.6%;	Pred. NO. 1.9;		
Matches	79;	Conservative 60;	Mismatches 120;	Indels 145; Gaps 16

Qy	135	SIHPELLKKYIEATQWVIGL-ATPLSPVDLVSNNQTAVSL-----PSPSHVLEA 184
Db	45	SIATFVNFCERFSY-YGMKAVILLYFLYFHWNEDTSTSYHAFSSICYETP--ILGA 101
Qy	185	QLNDS-KGEVLSSV--QMATTTEVLSLG-----OKLHAFAGLSVLLGTGLFM 231
Db	102	AADSWSLGKFTIIVLSLVVLGHVKISGALPILGGQVHTVLSLIGLSIALGTG-- 158
Qy	232	KRFIPCGWKYENIYITQVFQPERHRHQRPCVLPINKERRKYV-----II 280
Db	159	--GIKPCAAAPGG-----DQFEK-----HABERTFYSFVLLSINAGSLI 197
Qy	281	PTFWPTPKERNLKG-----LFFLPILIHLCIWLFSA-----ADVLL 317
Db	198	STP-IITMLRGDVOCGEDCALAFGPGLLMVIALVVFGMSKIYNKPPPEGNIVAQVF 256
Qy	318	YRLFYSVSQOPQS-----LPGFVEHL 338
Db	257	KCIWFATSNRPNKNSGDIPKRHDWLDAEKYPKLINDVKALTRVLFIYLPMPFWALL 316
Qy	339	KLHGKQGOTDI IHDSFNISVPFBNCPIPKPFLLSETWPL----- 380
Db	317	DQGSRWTLQAIRNRNLGFFVLQPDQQVLNPPLLVLIFIPLEDFVIYRLVSKCGINFSS 376
Qy	381	---SVILLIWMGLLSILMQLKTVLSASFVPSVERKRIQYLH 421
Db	377	LRKMAVGMIACLAFAAARVEIKINEMAPAQPQGVFLQVLN 420

RESULT 6

A40812
Ca2+-transporting ATPase (EC 3.6.3.8) SERCA2a - chicken
N:Alternate names: calcium pump
C:Species: Gallus gallus (chicken)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 19-Apr-2002
C:Accession: A40812
R:Campbell, A.M.; Kessler, P.D.; Sagara, Y.; Inesi, G.; Fambrough, D.M.
J. Biol. Chem. 266, 16050-16055, 1991
A:Title: Nucleotide sequences of avian cardiac and brain SR/ER Ca (2+)-ATPases and function of the nucleotide-binding domain
A:Reference number: A40812; MUID:91340754; PMID:1831452
A:Accession: A40812
A:Molecule type: mRNA
A:Residues: 1-997 <CAM>
A:Cross-references: GB:M66385; NID:G212652; PIDN:AAA49066.1; PID:G212653
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C:Keywords: ATP; hydrolase
F:594-767/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 4.1%; Score 100; DB 2; Length 997;
Best Local Similarity 25.4%; Pred. No. 3.1;
Matches 45; Conservative 29; Mismatches 59; Indels 44; Gaps 11;

Qy 21 PRSP-----GWMDFIQHLGVCCILVALISVGLLSVAACWEL-----PSIIATAAASWIITC 69
Db 820 PRNPKRPLISGWL-FFRYLAIGCVGAAATVG---AAAWFFIADGGPRVTFYQUSHFLQC 875
Qy 70 -----VLNCCSKHARCFTLLVFLSCGLRREGNALIAAGTGVILGHVENIFHNF 118
Db 876 KEONPDFSGVDCVVFESFPYMTALSVLVIEMCNLSL-SENQSLMRMPWENIW--- 931
Qy 119 KGLLDGWTGNLRAKSFSGIHFPLLKKYIEAIQWIVGLATPLSVFDDLVSNQTLAVSL 175
Db 932 ---LVGAIC---LSMSLHLFIL--YVEPLPIIFQI-TPLNV-----TQWLMLVLKISL 974

RESULT 7
S04651
Ca2+-transporting ATPase (EC 3.6.3.8) 1, sarcoplasmic/endoplasmic reticulum - pig
N:Alternate names: Ca2+-transporting ATPase isoform 2a; calcium pump
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 19-Apr-2002
C:Accession: S04651; S15078
R:Eggermont, J.A.; Wuytack, F.; De Jaegere, S.; Nelles, L.; Casteels, R.
Biochem. J. 260, 757-761, 1989
A:Title: Evidence for two isoforms of the endoplasmic-reticulum Ca(2+) pump in pig smooth muscle
A:Reference number: S04651; MUID:89350878; PMID:2527496
A:Accession: S04651
A:Molecule type: mRNA
A:Residues: 1-997 <EGG>
A:Cross-references: EMBL:X15073; NID:G1920; PIDN:CAA33169.1; PID:G1921
A:Experimental source: stomach; smooth muscle; sarcoplasmic/endoplasmic reticulum
R:Eggermont, J.A.; Wuytack, F.; Casteels, R.
Biochim. Biophys. Acta 1088, 448-451, 1991
A:Title: Characterization of the 3' end of the pig sarcoplasmic/endoplasmic-reticulum Ca(2+) pump
A:Reference number: S15077; MUID:91198156; PMID:2015309
A:Accession: S15078
A:Molecule type: DNA
A:Residues: 841-997 <EGW>
A:Cross-references: EMBL:X53754
A:Experimental source: sarcoplasmic/endoplasmic reticulum
C:Genetics:
A:Gene: SERCA2a
A:Introns: 841/1; 869/3; 914/2; 953/3; 994/1
C:Function:
A:Description: catalyzes hydrolysis of one molecule ATP coupled to translocation of two molecules such as the sarcoplasmic or endoplasmic reticulum; Ca2+ pump
A:Pathway: oxidative phosphorylation
A:Note: membrane-bound enzyme; magnesium-dependent; P-type ATPase
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C:Keywords: alternative splicing; ATP; calcium transport; endoplasmic reticulum; hydrolysis
F:60-57/Domain: calcium binding #status predicted <CAI>
F:60-78/Domain: transmembrane #status predicted <TM01>
F:87-107/Domain: transmembrane #status predicted <TM02>

RESULT 9

S04652

N;Alternate names: Ca2+-transporting ATPase (EC 3.6.3.8) 2, sarcoplasmic/endoplasmic reticulum - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 28-Feb-1990 #sequence revision 28-Feb-1990 #text_change 19-Apr-2002
C;Accession: S04652; S10244; S15077
R;Eggermont, J.A.; Wuytack, F.; De Jaegere, S.; Nelles, L.; Casteels, R.
Biochem. J. 260, 757-761, 1989
A;Title: Evidence for two isoforms of the endoplasmic-reticulum Ca(2+) pump in pig smooth muscle
A;Reference number: S04651; MUID:89350878; PMID:2527496
A;Accession: S04652
A;Molecule type: mRNA
A;Residues: 1-1042 <EGG>
A;Cross-references: EMBL:X15074; NID:g1922; PIDN:CAA33170.1; PID:g1923
R;Experimental source: smooth muscle; sarcoplasmic/endoplasmic reticulum
R;Eggermont, J.A.; Wuytack, F.; Casteels, R.
Biochem. J. 266, 901-907, 1990
A;Title: Characterization of the mRNAs encoding the gene 2 sarcoplasmic/endoplasmic-reticulum
A;Reference number: S10244; MUID:90226293; PMID:2158303
A;Accession: S10244
A;Molecule type: mRNA
A;Residues: 989-1042 <EG2>
A;Cross-references: EMBL:X16951; NID:g1924; PIDN:CAA34825.1; PID:gi364247
R;Experimental source: sarcoplasmic/endoplasmic reticulum
R;Eggermont, J.A.; Wuytack, F.; Casteels, R.
Biochim. Biophys. Acta 1088, 448-451, 1991
A;Title: Characterisation of the 3' end of the pig sarcoplasmic/endoplasmic-reticulum C
A;Reference number: S15077; MUID:91198156; PMID:2015309
A;Accession: S15077
A;Molecule type: DNA
A;Residues: 841-1042 <EGW>
A;Cross-references: EMBL:X53754
R;Experimental source: sarcoplasmic/endoplasmic reticulum
C;Genetics:
A;Gene: SERCA2b
A;Introns: 841/1; 869/3; 914/2; 953/3
C;Function:
es such as the sarcoplasmic or endoplasmic reticulum; Ca2+ pump
A;Pathway: oxidative phosphorylation
A;Note: membrane-bound enzyme; magnesium-dependent; P-type ATPase
C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C;Keywords: alternative splicing; ATP; calcium transport; endoplasmic reticulum; hydrolysis
F;40-57/Domain: calcium binding #status predicted <CA1>
F;60-78/Domain: transmembrane #status predicted <TM01>
F;87-107/Domain: transmembrane #status predicted <TM02>
F;108-257/Domain: intracellular #status predicted <INT1>
F;111-131/Domain: calcium binding #status predicted <CA2>
F;132-238/Domain: transduction #status predicted <TSD>
F;238-356/Domain: calcium binding #status predicted <CA3>
F;263-279/Domain: transmembrane #status predicted <TM03>
F;298-316/Domain: transmembrane #status predicted <TM04>
F;308-759/Domain: intracellular #status predicted <INT2>
F;310-329/Domain: calcium binding #status predicted <CA4>
F;330-505/Domain: catalytic #status predicted <PHY>
F;506-680/Domain: ATP binding #status predicted <ATP>
F;594-767/Domain: ATPase nucleotide-binding domain homology <ATW>
F;680-737/Domain: hinge #status predicted <HNG>
F;738-759/Domain: calcium binding #status predicted <CA5>
F;762-783/Domain: transmembrane #status predicted <TM05>
F;787-808/Domain: transmembrane #status predicted <TM06>
F;836-856/Domain: transmembrane #status predicted <TM07>
F;893-912/Domain: transmembrane #status predicted <TM08>
F;930-949/Domain: transmembrane #status predicted <TM09>
F;958-979/Domain: transmembrane #status predicted <TM10>
F;1013-1029/Domain: transmembrane #status predicted <TM11>
F;351/Active site: Asp (aspartylphosphate intermediate) #status predicted
F;514/Binding site: ATP (Lys) #status predicted

Query Match 4.1% Score 100: DB 2: Length 1042:

[illegible]

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84638
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-690 <STO>
A;Cross-references: GB:AE002093; NID:g4572671; PIDN:AAD23886.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g24610
A;Map position: 2

Query Match 4.1%; Score 99.5; DB 2; Length 690;
Best Local Similarity 24.9%; Pred. No. 2.1;
Matches 76; Conservative 36; Mismatches 104; Indels 89; Gaps 17;

Qy 49 VAACWFLPSIIAAAASWIITC-----VLLCCSKHARCFILLVFLSCGL--REGNALIAAG 102
Db :
227 LGSAYLLSIERQATCWAKESKESVPLQC-----VTDFDCTLHRDRDNW--QN 276
Qy 103 TGIIVI--IGHVENI FNFKGLLDGTCNLRAKSFSIHFFLLKKYIEAIOWIYGLATPLSV 160
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
277 TTWFNSNCDFSNNIQFTFGIFADALTKNV-----VSSPFLEKYLYCL-W-FGL----- 322
Qy 161 PDDLVSWNQTAVLSLSPSHVLEAQLNDSKEVLISLVLYQMATTTEVLSLGOKLLAFAGL 220
Db :
323 -QNLSSYGQNL-----STSTSVMETFMFALVAIFGL 352
Qy 221 SL--VLLGTGLFMKRFPGCGWKYENIYTRQVFQDERERHQRPQCVLPFNKEERKYV 278
Db | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
353 VLFALLIGN--MQTYLSQITVBLEEWLKRR-----DTEENWGHR--LLPQNLREVRVF 403
Qy 279 IIPTFWPTPKERNKLGFLLP-----LIHLCIWV-----LFAAVDVLLY-----RLIF 322
Db :
404 VOYKWLATRGVDEETILHSLPADLRDIQRHLCDLVRVPFLFAQMDQDLDAICERLAS 463
Qy 323 SVSKQ 327
Db | : | |
464 SLSTQ 468

RESULT 12
A30594
Ca2+-transporting ATPase (EC 3.6.3.8), cardiac muscle - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 16-Jun-1989 #sequence_revision 16-Jun-1989 #text_change 19-Apr-2002
C;Accession: A30594
R;Komuro, I.; Kurabayashi, M.; Shibasaki, Y.; Takaku, F.; Yazaki, Y.
J. Clin. Invest. 83, 1102-1108, 1989
A;Title: Molecular cloning and characterization of a Ca(2+)-dependent adenosine
A;Reference number: A30343; MUID:89198060; PMID:2522936
A;Accession: A30594
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-510 <KOM>
A;Cross-references: GB:M25267; NID:G602485; PIDN:AAA57270.1; PID:G602486
C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C;Keywords: ATP; cardiac muscle; heart; hydrolase
F;107-280/Domain: ATPase nucleotide-binding domain homology <TN>

Query Match 4.0%; Score 99; DB 2; Length 510;
Best Local Similarity 25.4%; Pred. No. 1.6;
Matches 45; Conservative 28; Mismatches 60; Indels 44; Gaps 11;

Qy 21 PRSP-----GWMDFIQHLGVCCVALISVGLLSVAACWFL-----PSIIAAASWIITC 69
Db ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
333 PRNPKEPLISGM-LPPRYLAIGCVGAATVG---AAAWFWIAADGGPRVSPFYQLSHFLQC 388
Qy 70 -----VLLCCSKHARCFILLVFLSCGLREGNALIAAGTGVILGHVENIFHNFP 118
Db KEDNPDFGVCDAIFESPYPMTMALSVLTMIEMCNLSL--SENOSLRMPPEWNIW-- 444

QY 119 KGLDGMTCNLRAKSFTHFPLKKYIEAIOIYGLATPLSVFDDLVSNQTLAVSL 175
Db 445 ----LVGSIC-----LSMSLHLFLL--YVEPLPLIFQI--TFLN-----NLQWLMLKISL 487

RESULT 13
PWBSBC
Ca2+-transporting ATPase (EC 3.6.3.8), slow twitch skeletal muscle - rabbit
N:Alternate names: calcium pump, slow twitch skeletal muscle; calcium-transporting ATPase
C:Species: *Oryctolagus cuniculus* (domestic rabbit)
C:Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 19-Apr-2002
C:Accession: A01076
R:MacLennan, D.H.; Brandl, C.J.; Korczak, B.; Green, N.M.
Nature 316, 696-700, 1985
A:Title: Amino-acid sequence of a Ca(2+)-Mg(2+)-dependent ATPase from rabbit muscle. sarc
A:Reference number: A01076; MUID:85296300; PMID:2993904
A:Accession: A01076
A:Molecule type: mRNA
A:Residues: 1-997 <MAC>
A:Cross-references: GB:X02814; GB:J02682; GB:M15159; NID:g1468; PIDN:CAA26583.1; PID:g14
C:Comment: This magnesium-dependent, membrane-bound enzyme catalyzes the hydrolysis of A
which skeletal muscle sarcolemmal reticulum; the slow twitch enzyme is low in activity.
n.
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C:Keywords: alternative splicing; ATP; calcium transport; hydrolase; magnesium; phospho
F:40-57/Domain: calcium binding #status predicted <CA1>
F:60-78/Domain: transmembrane #status predicted <TM01>
F:87-107/Domain: transmembrane #status predicted <TM02>
F:107-131/Domain: calcium binding #status predicted <CA2>
F:132-238/Domain: transduction <TSD>
F:238-256/Domain: calcium binding #status predicted <CA3>
F:258-277/Domain: transmembrane #status predicted <TM03>
F:288-307/Domain: transmembrane #status predicted <TM04>
F:310-329/Domain: calcium binding #status predicted <CA4>
F:506-680/Domain: nucleotide binding #status predicted <NBD>
F:594-767/Domain: ATPase nucleotide-binding domain homology <ATN>
F:680-740/Domain: hinge <HNG>
F:738-759/Domain: calcium binding #status predicted <CA5>
F:760-781/Domain: transmembrane #status predicted <TM05>
F:788-808/Domain: transmembrane #status predicted <TM06>
F:831-853/Domain: transmembrane #status predicted <TM07>
F:896-915/Domain: transmembrane #status predicted <TM08>
F:927-947/Domain: transmembrane #status predicted <TM09>
F:960-981/Domain: transmembrane #status predicted <TM10>
F:351/Active site: Asp (aspartylphosphate intermediate) #status predicted
F:514/Binding site: ATP (lys) #status predicted

Query Match 4.0%; Score 99; DB 1; Length 997;
Best Local Similarity 25.4%; Pred. No. 3.7;
Matches 45; Conservative 28; Mismatches 60; Indels 44; Gaps 11;

QY 21 PRSP-----GWMDFIQHLGVCCVLVALISVGLLSVAACWFL-----PSIIAAAASWIITC 69
Db 820 PRNPKEPLISGWL--FFRYLAIGCVGAATVG---AAAWFIAADGGRVSFYQLSHFLQC 875

QY 70 -----VLLCCSKHARCFTLLVFLSCGLREGNALIAAGTGIVILGHVENIPHN 118
Db 876 KEDNPDFEGVDCAFESPYMTWALSVLVTIEMCNALNSL--SENQSLRMPWENIW--- 931

119 KGLDGMTCNLRAKSFTHFPLKKYIEAIOIYGLATPLSVFDDLVSNQTLAVSL 175
Db 932 ---LVGSIC-----LSMSLHLFLL--YVEPLPLIFQI--TFLN-----TQWLMLKISL 974

RESULT 14
B31981
Ca2+-transporting ATPase (EC 3.6.3.8) 2, renal - human
C:Species: *Homo sapiens* (man)
C:Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 19-Apr-2002
C:Accession: B31981
R:Lytton, J.; MacLennan, D.H.
J. Biol. Chem. 263, 15024-15031, 1988
A:Title: Molecular cloning of cDNAs from human kidney coding for two alternatively splic

A:Reference number: A92673; MUID:89008384; PMID:2844796
A:Accession: B31981
A:Molecule type: mRNA
A:Residues: 1-997 <LYT>
A:Cross-references: GB:M23115; GB:J04025; NID:g184102; PIDN:AAA53194.1; PID:g306851
C:Genetics:
A:Gene: GDB:ATP2A2; ATP2B; SERCA2
A:Cross-references: GDB:1119717; OMIM:108740
A:Map position: 12q23-12q24.1
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C:Keywords: alternative splicing; ATP; hydrolase; phosphoprotein; transmembrane protein
F:594-767/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 4.0%; Score 99; DB 2; Length 997;
Best Local Similarity 25.4%; Pred. No. 3.7;
Matches 45; Conservative 28; Mismatches 60; Indels 44; Gaps 11;

QY 21 PRSP-----GWMDFIQHLGVCCVLVALISVGLLSVAACWFL-----PSIIAAAASWIITC 69
Db 820 PRNPKEPLISGWL--FFRYLAIGCVGAATVG---AAAWFIAADGGRVSFYQLSHFLQC 875

QY 70 -----VLLCCSKHARCFTLLVFLSCGLREGNALIAAGTGIVILGHVENIPHN 118
Db 876 KEDNPDFEGVDCAFESPYMTWALSVLVTIEMCNALNSL--SENQSLRMPWENIW--- 931

119 KGLDGMTCNLRAKSFTHFPLKKYIEAIOIYGLATPLSVFDDLVSNQTLAVSL 175
Db 932 ---LVGSIC-----LSMSLHLFLL--YVEPLPLIFQI--TFLN-----TQWLMLKISL 974

RESULT 15
B31982
Ca2+-transporting ATPase (EC 3.6.3.8) RS8-17 - rat
C:Species: *Rattus norvegicus* (Norway rat)
C:Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 19-Apr-2002
C:Accession: B31982; S04269
R:Guteski-Hamblin, A.M.; Greeb, J.; Shull, G.E.
J. Biol. Chem. 263, 15032-15040, 1988
A:Title: A novel Ca(2+) pump expressed in brain, kidney, and stomach is encoded by an
ncoding Ca(2+) and other cation-transporting ATPases using an oligonucleotide probe der
A:Reference number: A92674; MUID:89008385; PMID:2844797
A:Accession: B31982
A:Molecule type: mRNA
A:Residues: 1-997 <GUN>
A:Cross-references: GB:J04023; NID:g203058; PIDN:AAA40786.1; PID:g203059
A:Experimental source: stomach
R:Compre, A.M.; de la Bastie, D.; Boheler, K.R.; Schwartz, K.
FEBS Lett. 249, 35-41, 1989
A:Title: Characterization and expression of the rat heart sarcolemmal reticulum Ca(2+)
A:Reference number: S04269; MUID:89252068; PMID:2542094
A:Accession: S04269
A:Molecule type: mRNA
A:Residues: 1-997 <LOM>
A:Cross-references: EMBL:X15635; NID:g57302; PIDN:CAA33645.1; PID:g57303
A:Experimental source: skeletal and cardiac muscle
A:Note: the authors translated the codon TGG for residue 272 as Thr, TGG for residue 28
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C:Keywords: alternative splicing; ATP; calcium transport; hydrolase; membrane protein;
F:594-767/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 4.0%; Score 99; DB 2; Length 997;
Best Local Similarity 25.4%; Pred. No. 3.7;
Matches 45; Conservative 28; Mismatches 60; Indels 44; Gaps 11;

QY 21 PRSP-----GWMDFIQHLGVCCVLVALISVGLLSVAACWFL-----PSIIAAAASWIITC 69
Db 820 PRNPKEPLISGWL--FFRYLAIGCVGAATVG---AAAWFIAADGGRVSFYQLSHFLQC 875

QY 70 -----VLLCCSKHARCFTLLVFLSCGLREGNALIAAGTGIVILGHVENIPHN 118
Db 876 KEDNPDFEGVDCAFESPYMTWALSVLVTIEMCNALNSL--SENQSLRMPWENIW--- 931

119 KGLDGMTCNLRAKSFTHFPLKKYIEAIOIYGLATPLSVFDDLVSNQTLAVSL 175

42 Db 932 ---LVGSIC-----LSMSLHFLIL--YVEPLPLIFQI-TPL-----NLTCQWLMVKISL 974

Search completed: June 21, 2003, 13:26:30
Job time : 45 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2003, 13:18:47 ; Search time 23 Seconds
(without alignments)
847.560 Million cell updates/sec

Title: US-09-713-098-2

Perfect score: 2448

Sequence: 1 MGWISGTDIFLSLWEIVS.....LPVLKMRKKQMDMASADKS 470

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	102	4.2	447	Y341 CHLMU	Q9pkx0 chlamydia m
2	101	4.1	381	DHB2_RAT	Q62730 rattus norv
3	100.5	4.1	729	PET2 HUMAN	Q16348 homo sapien
4	100	4.1	1041	ATA2_CHICK	Q03669 g sarcoplas
5	100	4.1	1042	ATA2_PIG	P11607 s sarcoplas
6	99	4.0	997	ATA2_CANFA	Q46674 c sarcoplas
7	99	4.0	997	ATA2_FELCA	Q00779 f sarcoplas
8	99	4.0	1042	ATA2_HUMAN	P16615 h sarcoplas
9	99	4.0	1042	ATA2_RABIT	P20647 o sarcoplas
10	99	4.0	1043	ATA2_RAT	P11507 r sarcoplas
11	99	4.0	1044	ATA2_MOUSE	O55143 m sarcoplas
12	96	3.9	451	Y069_CHLTR	O84072 chlamydia t
13	96	3.9	500	YUEW_ECOLI	P39282 escherichia
14	94	3.8	729	PET2_RABIT	P46029 oryctolagus
15	93.5	3.8	622	VAI1_TREPA	O83444 treponema p
16	93.5	3.8	702	VATE_ENTEFA	Q93a44 enterococcu
17	93	3.8	543	EDS5_ARATH	Q945f0 arabidopsis
18	92	3.8	858	CVAG_DICDI	Q03101 dictyosteli
19	91	3.7	468	NAH_SCHPO	P36606 schizosacch
20	91	3.7	470	STZ3_YEAST	P06783 saccharomyc
21	91	3.7	515	NU2M_BETVU	P15688 beta vulgar
22	91	3.7	638	NU5M_PARLI	P12776 paracentrot
23	91	3.7	1581	ACC8_CRICR	Q09427 cricetus cr
24	90.5	3.7	282	UPK_MYCTU	O06239 mycobacteri
25	90.5	3.7	424	ER24_SCHPO	Q09195 schizosacch
26	90.5	3.7	1418	CE11_CABEL	P34641 caenorhabdi
27	90	3.7	706	Y006_RICPR	Q9zed6 rickettsia
28	90	3.7	2327	CCAB_MOUSE	O55017 mus musculu
29	89.5	3.7	369	TA2R_HUMAN	P21731 homo sapien
30	89.5	3.7	399	Y507_PASMU	Q9cnc5 pasteurella
31	89.5	3.7	463	NARK_ECOLI	P10903 escherichia
32	89.5	3.7	471	NORM_CAUCR	P58163 caulobacter
33	89.5	3.7	566	SYRD_PSESY	P33951 pseudomonas

34	89.5	3.7	674	1	NUOL_NEIMB	Q9klb0 neisseria m
35	89	3.6	282	1	FOCE_ECOLI	P77733 escherichia
36	89	3.6	311	1	O1D4_HUMAN	P47884 homo sapien
37	89	3.6	312	1	O1D2_HUMAN	P34982 homo sapien
38	89	3.6	374	1	Y006_BORBU	O51039 borrelia bu
39	89	3.6	497	1	NU2M_CHOCR	P48903 chondrus cr
40	89	3.6	563	1	Y1M0_YEAST	P40475 saccharomyc
41	89	3.6	2336	1	CCAB_RAT	Q02294 rattus norv
42	88.5	3.6	674	1	NUOL_NEIMA	Q9jx52 neisseria m
43	88.5	3.6	1203	1	YT41_CABEL	Q11069 caenorhabdi
44	88.5	3.6	2339	1	CCAB_HUMAN	Q00975 homo sapien
45	88	3.6	453	1	CCKE_XENLA	P70031 xenopus lae

ALIGNMENTS

RESULT 1
Y341_CHLMU STANDARD; PRT; 447 AA.
AC Q9PKX0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable metal transport system membrane protein TC0341.
GN TC0341.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MoPn / Nigg;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., Bass S., White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bowdson R., Linner K., Weidman J., Khouri H., Craven B., Bowman C., Szalberg S.L., Gwin M., Nelson W., DeBoy J., McClarty J., Kolonay J., McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -!- FUNCTION: PART OF AN ATP-DRIVEN TRANSPORT SYSTEM
CC TC0338/TC0339/TC0341/TC0342 FOR A METAL.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (Probable).
CC -!- SIMILARITY: BELONGS TO THE ABC-3 SUBFAMILY OF INTEGRAL MEMBRANE PROTEINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AE002301; AAF73548.1; --
TIGR; TC0341; --
InterPro; IPR001626; ABCtransprtr3.
DR InterPro; IPR001367; HTH_DTXR.
DR Pfam; PF00950; ABC-3; 1.
DR Pfam; PF02742; Fe dep repr C; 1.
DR SMART; SM00529; HTH_DTXR; 1.
KW Hypothetical protein; Transmembrane; Transmembrane; Inner membrane; Complete proteome.
FT TRANSMEM 14 34 POTENTIAL.
FT TRANSMEM 38 58 POTENTIAL.
FT TRANSMEM 69 89 POTENTIAL.
FT TRANSMEM 100 120 POTENTIAL.
FT TRANSMEM 152 172 POTENTIAL.
FT TRANSMEM 192 212 POTENTIAL.
FT TRANSMEM 233 253 POTENTIAL.
FT TRANSMEM 270 290 POTENTIAL.

SQ SEQUENCE 447 AA; 50573 MW; BA65E5CB3796452E CRC64;
 Query Match 4.2%; Score 102; DB 1; Length 447;
 Best Local Similarity 19.0%; Pred. No. 1.2;
 Matches 74; Conservative 59; Mismatches 111; Indels 146; Gaps 19;
 QY 10 IFLS-----LW-EYVSPRSGWDFIOHGLVCCVALISVGLLSVAACWFL 55
 DB 10 IFLSFLAVSLICMTALWGTLVRRPLLSLSH-----ACVPGLLIGALLSYKVPFLS 66
 QY 56 PS-----IAAASWIITCVLLCCS-----KHARCFILLVFLSCGLREGNALLIAA 101
 DB 67 DSLWIIILCGCAASILGC--LCIAFLEKKLAMHKOSALCILVSPF-----110
 QY 102 GTGIVLGHVENIFHNFKLLDGMTCNLRAKSFHPLKXYIEAIOIWIYGLATPLSPV 161
 DB 111 GVGIVLSVYKD-----SC-----PLLYNKINA--YLYGQATLGYA 145
 QY 162 DDLVS-----WNQTLAVSLSPSHVLEAQLNDSKGEVLSVLYQWATTTEVLS 208
 DB 146 EARLALIIFCSAFVLMWYRQISVAIFDREPAYSCGLRTRTAEMIVLIFSLIVSGVR 205
 QY 209 SLGCKLLA--FAG-----LSLVLLGTGLFMKRLPGCG--WKYENIYITRQFVQ 253
 DB 206 SVGILLISAMFVAPLSARQLSKSLILSIFG-----GICGALGCYFSVAFTCSQSI- 260
 QY 254 FDERERHOORPCVLPINKEERKYVIPTFWTPKRNKLGFLFFLPILHLICI----- 306
 DB 261 -----TEGKATVILPT-----GPLVWFFAGVLFCLIFSKWTKG 295
 QY 307 WV-----LFAADVILLYRLIFSVSKQ 327
 DB 296 WITRYIRKFLFRSDEEHLKIFWYLOEQ 325

RESULT 2

DHB2 RAT
 ID DHB2 RAT STANDARD; PRT; 381 AA.
 AC Q62730;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Estradiol 17 beta-dehydrogenase 2 (EC 1.1.1.62) (17-beta-HSD 2)
 DE (17-beta-hydroxysteroid dehydrogenase 2).
 GN HSD17B2 OR EDH17B2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Placenta;
 RX MEDLINE=96198748; PubMed=8612487;
 RA Akinola L.A., Poutanen M., Viikko R.;
 RT "Cloning of rat 17 beta-hydroxysteroid dehydrogenase type 2 and
 RT characterization of tissue distribution and catalytic activity of rat
 RT type 1 and type 2 enzymes."
 RL Endocrinology 137:1572-1579 (1996).
 CC -!- FUNCTION: CAPABLE OF CATALYZING THE INTERCONVERSION OF
 CC TESTOSTERONE AND ANDROSTENEDIONE, AS WELL AS ESTRADIOL AND
 CC ESTRONE. ALSO HAS 20-ALPHA-HSD ACTIVITY. USES NADH WHILE EDH17B3
 CC USES NADPH (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: Estradiol-17-beta + NAD(P)(+) = estrone +
 CC NAD(P)H.
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
 CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN PLACENTA.
 CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by 'non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X91234; CAA62617.1; -;
 DR InterPro; IPR002198; ADH_short.
 DR Pfam; PF00106; adh_short; 1.
 DR PRINTS; PR00080; SDRFAMILY.
 DR PROSITE; PS00061; ADH_SHORT; 1.
 KW Steroid biosynthesis; Oxidoreductase; NAD; Multigene family;
 FT TRANSMEM 4 24 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
 FT NP_BIND 83 112 NAD (BY SIMILARITY).
 FT ACT_SITE 233 BY SIMILARITY.
 SQ SEQUENCE 381 AA; 41967 MW; 8E1D08DF345DE136 CRC64;
 Query Match 4.1%; Score 101; DB 1; Length 381;
 Best Local Similarity 21.0%; Pred. No. 1.2;
 Matches 78; Conservative 51; Mismatches 133; Indels 110; Gaps 22;
 QY 37 CLVALISVGLLSVAACW-----FLPSITAAASWIITCVL---LCCS---KHARCFILLVF 86
 DB 12 CLTATAVLGMLLCKAWSSQQLRSQVCLAGLWGGACLLSLSLCSFLLSVSCFFLLYV 71
 QY 87 LSCG---LREGNALLTAAGT---GIVILGHVENI-FHNFKGLLD-----GMCNVL 129
 DB 72 SSSDQDLLPVDQKAVLVTGADSGFHALAKHLKGLGTFVAGVLDKEGPGAELRKNCS 131
 QY 130 RAKSF-----SIHFLLKXYIEAIOW-----IYGLATPLSVDFDLVS 166
 DB 132 RLSQLQMDVTKPQIKDVHSEVAEKIQDKGLMANNAGVLHPIDGELIPMTVI--- 186
 QY 167 WNQTLAVSLFSPSHVLEA---QLNDSKGEVLSV---YQW---ATTTEVLSSIG- 211
 DB 187 -KCMVNFEGAVEVTKVFLPLLRKSKGLRVNVSVMGAMIPFQWVAAYASTKAISMESA 245
 QY 212 --OKLLAFAGLSLVLLGTGLFMKRLG-PCGWKYENIYITRQFVQ-----FDERERHOOR 263
 DB 246 VIRQELAKWGVKVTIHPGGFQTNIVGSDQSDWDMKEILDHFSKETQENYQGEYVHTQK 305
 QY 264 PCVLPINKEERKYVIPTFWTPKE-----RKNLGLFPLPILHLIC-----IWLFA 311
 DB 306 -LALPYRMESN-----PDITPVLRIQHAICAKNPSFY-----CSGRMTYLIWICFA 352
 QY 312 A-----VDYLL 317
 DB 353 AYSPISLLDYIL 364

RESULT 3

PET2 HUMAN
 ID PET2 HUMAN STANDARD; PRT; 729 AA.
 AC Q16378;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Oligopeptide transporter, kidney isoform (peptide transporter 2)
 DE (Kidney H+/peptide cotransporter) (Solute carrier family 15, member
 DE 2).
 DE SLC15A2 OR PEPT2.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=95275926; PubMed=7756356;
 RA Liu W., Liang R., Ramamoorthy S., Fei Y.J., Ganapathy M.E.,
 RA Rediger M.A., Ganapathy V., Leibach F.H.;
 RT "Molecular cloning of PEPT 2, a new member of the H+/peptide
 RT cotransporter family, from human kidney."
 RL Biochim. Biophys. Acta 1235:461-466 (1995).

CC -!- FUNCTION: PROTON-COUPLED INTAKE OF OLIGOPEPTIDES OF 2 TO 4
CC AMINO ACIDS WITH A PREFERENCE FOR DIPEPTIDES.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE PTR2 FAMILY OF TRANSPORTERS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; S78203; AAB34388.1; -.
DR Genew; HGNC:10921; SLC15A2.
DR MIM; 602339; -.
DR InterPro; IPR00109; PTR2.
DR InterPro; IPR004768; Pep_H_symport.
DR Pfam; PF00854; PTR2; 2.
DR TIGRPFAMS; TIGR00926; 2A1704; 1.
DR PROSITE; PS01022; PTR2_1; 1.
DR PROSITE; PS01023; PTR2_2; 1.
KW Peptide transport; Transport; Transmembrane; Symport; Glycoprotein.
FT TRANSMEM 58 78
FT TRANSMEM 88 108
FT TRANSMEM 115 135
FT TRANSMEM 140 160
FT TRANSMEM 184 204
FT TRANSMEM 218 238
FT TRANSMEM 296 316
FT TRANSMEM 344 364
FT TRANSMEM 381 401
FT TRANSMEM 568 588
FT TRANSMEM 612 632
FT TRANSMEM 644 664
FT TRANSMEM 675 695
FT CARBOHYD 7 7
FT CARBOHYD 269 269
FT CARBOHYD 373 373
FT CARBOHYD 435 435
FT CARBOHYD 472 472
FT CARBOHYD 528 528
FT CARBOHYD 567 567
SQ SEQUENCE 729 AA; 81940 MW; F046073D27C063D3 CRC64;

Query Match 4.1%; Score 100.5; DB 1; Length 729;
Best Local Similarity 19.6%; Pred. No. 2.8; Mismatches 120; Indels 145; Gaps 16;
Matches 79; Conservative 60;
QY 135 SIHFPLKAKYIAQWIXGL-ATPLSVFDDLVSNWOTLAVSL-----FSPSHVLEA 184
DB 45 SIATFVNEPCEFSY-YGMKAVLLYFLYFLHWNEDTSTSIYHAFSSLCYFTP--ILGA 101
QY 185 QUNDS-KGEVLSVLY-QWATTEVLSLG-----QKLAPAGLSVLLGTGLFM 231
DB 102 ATADSLGKFKXTIIVLSLVVYLVGHVKSIGLALPILGGQVWVTVLSLIGLSLALGTG--- 158
QY 232 KFLGPGCKWKNYITQFQVDFDRHRHQPCVPLNKEERKVV-----II 280
DB 159 --GKPCVAARFGG-----DQFEK-----HAERTYFVSFVLSINAGSLI 197
QY 281 PTFWPTPKERKNLG-----LFFLPILHLCTIWLPLA-----AVDYLL 317
DB 198 STF-ITPMLRGVQCGEDCVALAFVPGLLMVIALVWFAMGSKYINKPPPGENVAQVF 256
QY 318 YRLIFSVSQKQFS-----LPGFEVHL 338
DB 257 KCIWFAISNFRNRSGLDKPRHDLWDAAEKYKPKQLIMDKALTRVFLYIPLPFWALL 316
QY 339 KLHGKQGTQDIHDSSFNISVFEPNCPKPKELLSETWVPL----- 380
DB 317 DQGSRTWLQAIRNRNLGFFVLQPDQMVLPPLLVLPFLPFDVIVRLVSKCGINFS 376

QY 381 ---SVILLIVMLGLSSILMOLKILVSASFYSVERKRIOYLH 421
DB 377 LRMAVGMILACLAFVAARVEIKINEMAPQGPQEVFLQVLN 420
RESULT 4
AT2A2_CHICK STANDARD; PRT; 1041 AA.
ID AT2A2_CHICK AC Q03669;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sarcoplasmic/endoplasmic reticulum calcium ATPase 2 (EC 3.6.3.8)
DE (Calcium pump 2) (SERCA2) (SR Ca(2+)-ATPase 2) (Calcium-transporting
DE ATPase sarcoplasmic reticulum type, slow twitch skeletal muscle
DE isoform) (Endoplasmic reticulum class 1/2 Ca(2+) ATPase).
GN ATP2A2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]_TaxID=9031;
RP SEQUENCE FROM N.A. (ISOFORMS SERCA2A AND SERCA2B).
RC STRAIN=White leghorn; TISSUE=Heart, and Brain;
RX MEDLINE=91340754; PubMed=1831452;
RA Campbell A.M., Kessler P.D., Sagara Y., Fambrough D.M.;
RT "Nucleotide sequences of avian cardiac and brain SR/ER Ca(2+)-ATPases
RT and functional comparisons with fast twitch Ca(2+)-ATPase. Calcium
RT affinities and inhibitor effects";
RL J. Biol. Chem. 266:16050-16055(1991).
CC -!- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE HYDROLYSIS
CC OF ATP COUPLED WITH THE TRANSLOCATION OF CALCIUM FROM THE CYTOSOL
CC TO THE SARCOPLASMIC RETICULUM LUMEN. ISOFORM SERCA2A IS INVOLVED
CC IN THE REGULATION OF THE CONTRACTION/RELAXATION CYCLE (BY
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + Ca(2+) (in) = ADP + phosphate +
CC Ca(2+) (out).
CC -!- ENZYME REGULATION: REVERSIBLY INHIBITED BY PHOSPHOLAMBAN (PLN) AT
CC LOW CALCIUM CONCENTRATIONS. DEPHOSPHORYLATED PLN DECREASES THE
CC APPARENT AFFINITY OF THE ATPASE FOR CALCIUM. THIS INHIBITION IS
CC REGULATED BY THE PHOSPHORYLATION OF PLN (BY SIMILARITY).
CC -!- SUBUNIT: ASSOCIATED WITH PHOSPHOLAMBAN (PLN) (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. SARCOPLASMIC AND
CC ENDOPLASMIC RETICULUM.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; SERCA2A/ATP2A2A AND
CC SERCA2B/ATP2A2B (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING
CC IN A TISSUE-SPECIFIC MANNER.
CC -!- TISSUE SPECIFICITY: ONLY ISOFORM SERCA2A IS DETECTED IN HEART,
CC WHILE BOTH ISOFORMS ARE EXPRESSED IN BRAIN, WITH SERCA2B BEING THE
CC PREDOMINANT FORM.
CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
CC (E1-E2 ATPASES). SUBFAMILY IIA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M66385; AAA49066.1; -.
DR HSP; P04191; LEUL.
DR InterPro; IPR001757; ATPase_E1-E2.
DR InterPro; IPR004014; Cation ATPase.
DR InterPro; IPR001454; Hlgase/hydrlase.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00689; Cation ATPase; 1.
DR Pfam; PF00690; Cation ATPase; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR PRINTS; PR00119; CATATPASE.
DR PROSITE; PS00154; ATPASE_E1_E2; 1.


```

FT TRANSMEM 90 110 2 (BY SIMILARITY).
FT DOMAIN 111 253 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 254 273 3 (BY SIMILARITY).
FT DOMAIN 274 295 LUMENAL (BY SIMILARITY).
FT TRANSMEM 296 313 4 (BY SIMILARITY).
FT DOMAIN 314 756 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 757 776 5 (BY SIMILARITY).
FT DOMAIN 777 786 LUMENAL (BY SIMILARITY).
FT TRANSMEM 787 807 6 (BY SIMILARITY).
FT DOMAIN 808 827 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 828 850 7 (BY SIMILARITY).
FT DOMAIN 851 896 LUMENAL (BY SIMILARITY).
FT TRANSMEM 897 916 8 (BY SIMILARITY).
FT DOMAIN 917 929 9 (BY SIMILARITY).
FT TRANSMEM 930 948 10 (BY SIMILARITY).
FT DOMAIN 949 963 11 (BY SIMILARITY).
FT TRANSMEM 964 984 12 (BY SIMILARITY).
FT DOMAIN 985 1042 13 (BY SIMILARITY).
FT TRANSMEM 1043 1062 14 (BY SIMILARITY).
FT DOMAIN 1063 1082 15 (BY SIMILARITY).
FT TRANSMEM 1083 1102 16 (BY SIMILARITY).
FT DOMAIN 1103 1122 17 (BY SIMILARITY).
FT TRANSMEM 1123 1142 18 (BY SIMILARITY).
FT DOMAIN 1143 1162 19 (BY SIMILARITY).
FT TRANSMEM 1163 1182 20 (BY SIMILARITY).
FT DOMAIN 1183 1202 21 (BY SIMILARITY).
FT TRANSMEM 1203 1222 22 (BY SIMILARITY).
FT DOMAIN 1223 1242 23 (BY SIMILARITY).
FT TRANSMEM 1243 1262 24 (BY SIMILARITY).
FT DOMAIN 1263 1282 25 (BY SIMILARITY).
FT TRANSMEM 1283 1302 26 (BY SIMILARITY).
FT DOMAIN 1303 1322 27 (BY SIMILARITY).
FT TRANSMEM 1323 1342 28 (BY SIMILARITY).
FT DOMAIN 1343 1362 29 (BY SIMILARITY).
FT TRANSMEM 1363 1382 30 (BY SIMILARITY).
FT DOMAIN 1383 1402 31 (BY SIMILARITY).
FT TRANSMEM 1403 1422 32 (BY SIMILARITY).
FT DOMAIN 1423 1442 33 (BY SIMILARITY).
FT TRANSMEM 1443 1462 34 (BY SIMILARITY).
FT DOMAIN 1463 1482 35 (BY SIMILARITY).
FT TRANSMEM 1483 1502 36 (BY SIMILARITY).
FT DOMAIN 1503 1522 37 (BY SIMILARITY).
FT TRANSMEM 1523 1542 38 (BY SIMILARITY).
FT DOMAIN 1543 1562 39 (BY SIMILARITY).
FT TRANSMEM 1563 1582 40 (BY SIMILARITY).
FT DOMAIN 1583 1602 41 (BY SIMILARITY).
FT TRANSMEM 1603 1622 42 (BY SIMILARITY).
FT DOMAIN 1623 1642 43 (BY SIMILARITY).
FT TRANSMEM 1643 1662 44 (BY SIMILARITY).
FT DOMAIN 1663 1682 45 (BY SIMILARITY).
FT TRANSMEM 1683 1702 46 (BY SIMILARITY).
FT DOMAIN 1703 1722 47 (BY SIMILARITY).
FT TRANSMEM 1723 1742 48 (BY SIMILARITY).
FT DOMAIN 1743 1762 49 (BY SIMILARITY).
FT TRANSMEM 1763 1782 50 (BY SIMILARITY).
FT DOMAIN 1783 1802 51 (BY SIMILARITY).
FT TRANSMEM 1803 1822 52 (BY SIMILARITY).
FT DOMAIN 1823 1842 53 (BY SIMILARITY).
FT TRANSMEM 1843 1862 54 (BY SIMILARITY).
FT DOMAIN 1863 1882 55 (BY SIMILARITY).
FT TRANSMEM 1883 1902 56 (BY SIMILARITY).
FT DOMAIN 1903 1922 57 (BY SIMILARITY).
FT TRANSMEM 1923 1942 58 (BY SIMILARITY).
FT DOMAIN 1943 1962 59 (BY SIMILARITY).
FT TRANSMEM 1963 1982 60 (BY SIMILARITY).
FT DOMAIN 1983 2002 61 (BY SIMILARITY).
FT TRANSMEM 2003 2022 62 (BY SIMILARITY).
FT DOMAIN 2023 2042 63 (BY SIMILARITY).
FT TRANSMEM 2043 2062 64 (BY SIMILARITY).
FT DOMAIN 2063 2082 65 (BY SIMILARITY).
FT TRANSMEM 2083 2102 66 (BY SIMILARITY).
FT DOMAIN 2103 2122 67 (BY SIMILARITY).
FT TRANSMEM 2123 2142 68 (BY SIMILARITY).
FT DOMAIN 2143 2162 69 (BY SIMILARITY).
FT TRANSMEM 2163 2182 70 (BY SIMILARITY).
FT DOMAIN 2183 2202 71 (BY SIMILARITY).
FT TRANSMEM 2203 2222 72 (BY SIMILARITY).
FT DOMAIN 2223 2242 73 (BY SIMILARITY).
FT TRANSMEM 2243 2262 74 (BY SIMILARITY).
FT DOMAIN 2263 2282 75 (BY SIMILARITY).
FT TRANSMEM 2283 2302 76 (BY SIMILARITY).
FT DOMAIN 2303 2322 77 (BY SIMILARITY).
FT TRANSMEM 2323 2342 78 (BY SIMILARITY).
FT DOMAIN 2343 2362 79 (BY SIMILARITY).
FT TRANSMEM 2363 2382 80 (BY SIMILARITY).
FT DOMAIN 2383 2402 81 (BY SIMILARITY).
FT TRANSMEM 2403 2422 82 (BY SIMILARITY).
FT DOMAIN 2423 2442 83 (BY SIMILARITY).
FT TRANSMEM 2443 2462 84 (BY SIMILARITY).
FT DOMAIN 2463 2482 85 (BY SIMILARITY).
FT TRANSMEM 2483 2502 86 (BY SIMILARITY).
FT DOMAIN 2503 2522 87 (BY SIMILARITY).
FT TRANSMEM 2523 2542 88 (BY SIMILARITY).
FT DOMAIN 2543 2562 89 (BY SIMILARITY).
FT TRANSMEM 2563 2582 90 (BY SIMILARITY).
FT DOMAIN 2583 2602 91 (BY SIMILARITY).
FT TRANSMEM 2603 2622 92 (BY SIMILARITY).
FT DOMAIN 2623 2642 93 (BY SIMILARITY).
FT TRANSMEM 2643 2662 94 (BY SIMILARITY).
FT DOMAIN 2663 2682 95 (BY SIMILARITY).
FT TRANSMEM 2683 2702 96 (BY SIMILARITY).
FT DOMAIN 2703 2722 97 (BY SIMILARITY).
FT TRANSMEM 2723 2742 98 (BY SIMILARITY).
FT DOMAIN 2743 2762 99 (BY SIMILARITY).
FT TRANSMEM 2763 2782 100 (BY SIMILARITY).
FT DOMAIN 2783 2802 101 (BY SIMILARITY).
FT TRANSMEM 2803 2822 102 (BY SIMILARITY).
FT DOMAIN 2823 2842 103 (BY SIMILARITY).
FT TRANSMEM 2843 2862 104 (BY SIMILARITY).
FT DOMAIN 2863 2882 105 (BY SIMILARITY).
FT TRANSMEM 2883 2902 106 (BY SIMILARITY).
FT DOMAIN 2903 2922 107 (BY SIMILARITY).
FT TRANSMEM 2923 2942 108 (BY SIMILARITY).
FT DOMAIN 2943 2962 109 (BY SIMILARITY).
FT TRANSMEM 2963 2982 110 (BY SIMILARITY).
FT DOMAIN 2983 3002 111 (BY SIMILARITY).
FT TRANSMEM 3003 3022 112 (BY SIMILARITY).
FT DOMAIN 3023 3042 113 (BY SIMILARITY).
FT TRANSMEM 3043 3062 114 (BY SIMILARITY).
FT DOMAIN 3063 3082 115 (BY SIMILARITY).
FT TRANSMEM 3083 3102 116 (BY SIMILARITY).
FT DOMAIN 3103 3122 117 (BY SIMILARITY).
FT TRANSMEM 3123 3142 118 (BY SIMILARITY).
FT DOMAIN 3143 3162 119 (BY SIMILARITY).
FT TRANSMEM 3163 3182 120 (BY SIMILARITY).
FT DOMAIN 3183 3202 121 (BY SIMILARITY).
FT TRANSMEM 3203 3222 122 (BY SIMILARITY).
FT DOMAIN 3223 3242 123 (BY SIMILARITY).
FT TRANSMEM 3243 3262 124 (BY SIMILARITY).
FT DOMAIN 3263 3282 125 (BY SIMILARITY).
FT TRANSMEM 3283 3302 126 (BY SIMILARITY).
FT DOMAIN 3303 3322 127 (BY SIMILARITY).
FT TRANSMEM 3323 3342 128 (BY SIMILARITY).
FT DOMAIN 3343 3362 129 (BY SIMILARITY).
FT TRANSMEM 3363 3382 130 (BY SIMILARITY).
FT DOMAIN 3383 3402 131 (BY SIMILARITY).
FT TRANSMEM 3403 3422 132 (BY SIMILARITY).
FT DOMAIN 3423 3442 133 (BY SIMILARITY).
FT TRANSMEM 3443 3462 134 (BY SIMILARITY).
FT DOMAIN 3463 3482 135 (BY SIMILARITY).
FT TRANSMEM 3483 3502 136 (BY SIMILARITY).
FT DOMAIN 3503 3522 137 (BY SIMILARITY).
FT TRANSMEM 3523 3542 138 (BY SIMILARITY).
FT DOMAIN 3543 3562 139 (BY SIMILARITY).
FT TRANSMEM 3563 3582 140 (BY SIMILARITY).
FT DOMAIN 3583 3602 141 (BY SIMILARITY).
FT TRANSMEM 3603 3622 142 (BY SIMILARITY).
FT DOMAIN 3623 3642 143 (BY SIMILARITY).
FT TRANSMEM 3643 3662 144 (BY SIMILARITY).
FT DOMAIN 3663 3682 145 (BY SIMILARITY).
FT TRANSMEM 3683 3702 146 (BY SIMILARITY).
FT DOMAIN 3703 3722 147 (BY SIMILARITY).
FT TRANSMEM 3723 3742 148 (BY SIMILARITY).
FT DOMAIN 3743 3762 149 (BY SIMILARITY).
FT TRANSMEM 3763 3782 150 (BY SIMILARITY).
FT DOMAIN 3783 3802 151 (BY SIMILARITY).
FT TRANSMEM 3803 3822 152 (BY SIMILARITY).
FT DOMAIN 3823 3842 153 (BY SIMILARITY).
FT TRANSMEM 3843 3862 154 (BY SIMILARITY).
FT DOMAIN 3863 3882 155 (BY SIMILARITY).
FT TRANSMEM 3883 3902 156 (BY SIMILARITY).
FT DOMAIN 3903 3922 157 (BY SIMILARITY).
FT TRANSMEM 3923 3942 158 (BY SIMILARITY).
FT DOMAIN 3943 3962 159 (BY SIMILARITY).
FT TRANSMEM 3963 3982 160 (BY SIMILARITY).
FT DOMAIN 3983 4002 161 (BY SIMILARITY).
FT TRANSMEM 4003 4022 162 (BY SIMILARITY).
FT DOMAIN 4023 4042 163 (BY SIMILARITY).
FT TRANSMEM 4043 4062 164 (BY SIMILARITY).
FT DOMAIN 4063 4082 165 (BY SIMILARITY).
FT TRANSMEM 4083 4102 166 (BY SIMILARITY).
FT DOMAIN 4103 4122 167 (BY SIMILARITY).
FT TRANSMEM 4123 4142 168 (BY SIMILARITY).
FT DOMAIN 4143 4162 169 (BY SIMILARITY).
FT TRANSMEM 4163 4182 170 (BY SIMILARITY).
FT DOMAIN 4183 4202 171 (BY SIMILARITY).
FT TRANSMEM 4203 4222 172 (BY SIMILARITY).
FT DOMAIN 4223 4242 173 (BY SIMILARITY).
FT TRANSMEM 4243 4262 174 (BY SIMILARITY).
FT DOMAIN 4263 4282 175 (BY SIMILARITY).
FT TRANSMEM 4283 4302 176 (BY SIMILARITY).
FT DOMAIN 4303 4322 177 (BY SIMILARITY).
FT TRANSMEM 4323 4342 178 (BY SIMILARITY).
FT DOMAIN 4343 4362 179 (BY SIMILARITY).
FT TRANSMEM 4363 4382 180 (BY SIMILARITY).
FT DOMAIN 4383 4402 181 (BY SIMILARITY).
FT TRANSMEM 4403 4422 182 (BY SIMILARITY).
FT DOMAIN 4423 4442 183 (BY SIMILARITY).
FT TRANSMEM 4443 4462 184 (BY SIMILARITY).
FT DOMAIN 4463 4482 185 (BY SIMILARITY).
FT TRANSMEM 4483 4502 186 (BY SIMILARITY).
FT DOMAIN 4503 4522 187 (BY SIMILARITY).
FT TRANSMEM 4523 4542 188 (BY SIMILARITY).
FT DOMAIN 4543 4562 189 (BY SIMILARITY).
FT TRANSMEM 4563 4582 190 (BY SIMILARITY).
FT DOMAIN 4583 4602 191 (BY SIMILARITY).
FT TRANSMEM 4603 4622 192 (BY SIMILARITY).
FT DOMAIN 4623 4642 193 (BY SIMILARITY).
FT TRANSMEM 4643 4662 194 (BY SIMILARITY).
FT DOMAIN 4663 4682 195 (BY SIMILARITY).
FT TRANSMEM 4683 4702 196 (BY SIMILARITY).
FT DOMAIN 4703 4722 197 (BY SIMILARITY).
FT TRANSMEM 4723 4742 198 (BY SIMILARITY).
FT DOMAIN 4743 4762 199 (BY SIMILARITY).
FT TRANSMEM 4763 4782 200 (BY SIMILARITY).
FT DOMAIN 4783 4802 201 (BY SIMILARITY).
FT TRANSMEM 4803 4822 202 (BY SIMILARITY).
FT DOMAIN 4823 4842 203 (BY SIMILARITY).
FT TRANSMEM 4843 4862 204 (BY SIMILARITY).
FT DOMAIN 4863 4882 205 (BY SIMILARITY).
FT TRANSMEM 4883 4902 206 (BY SIMILARITY).
FT DOMAIN 4903 4922 207 (BY SIMILARITY).
FT TRANSMEM 4923 4942 208 (BY SIMILARITY).
FT DOMAIN 4943 4962 209 (BY SIMILARITY).
FT TRANSMEM 4963 4982 210 (BY SIMILARITY).
FT DOMAIN 4983 5002 211 (BY SIMILARITY).
FT TRANSMEM 5003 5022 212 (BY SIMILARITY).
FT DOMAIN 5023 5042 213 (BY SIMILARITY).
FT TRANSMEM 5043 5062 214 (BY SIMILARITY).
FT DOMAIN 5063 5082 215 (BY SIMILARITY).
FT TRANSMEM 5083 5102 216 (BY SIMILARITY).
FT DOMAIN 5103 5122 217 (BY SIMILARITY).
FT TRANSMEM 5123 5142 218 (BY SIMILARITY).
FT DOMAIN 5143 5162 219 (BY SIMILARITY).
FT TRANSMEM 5163 5182 220 (BY SIMILARITY).
FT DOMAIN 5183 5202 221 (BY SIMILARITY).
FT TRANSMEM 5203 5222 222 (BY SIMILARITY).
FT DOMAIN 5223 5242 223 (BY SIMILARITY).
FT TRANSMEM 5243 5262 224 (BY SIMILARITY).
FT DOMAIN 5263 5282 225 (BY SIMILARITY).
FT TRANSMEM 5283 5302 226 (BY SIMILARITY).
FT DOMAIN 5303 5322 227 (BY SIMILARITY).
FT TRANSMEM 5323 5342 228 (BY SIMILARITY).
FT DOMAIN 5343 5362 229 (BY SIMILARITY).
FT TRANSMEM 5363 5382 230 (BY SIMILARITY).
FT DOMAIN 5383 5402 231 (BY SIMILARITY).
FT TRANSMEM 5403 5422 232 (BY SIMILARITY).
FT DOMAIN 5423 5442 233 (BY SIMILARITY).
FT TRANSMEM 5443 5462 234 (BY SIMILARITY).
FT DOMAIN 5463 5482 235 (BY SIMILARITY).
FT TRANSMEM 5483 5502 236 (BY SIMILARITY).
FT DOMAIN 5503 5522 237 (BY SIMILARITY).
FT TRANSMEM 5523 5542 238 (BY SIMILARITY).
FT DOMAIN 5543 5562 239 (BY SIMILARITY).
FT TRANSMEM 5563 5582 240 (BY SIMILARITY).
FT DOMAIN 5583 5602 241 (BY SIMILARITY).
FT TRANSMEM 5603 5622 242 (BY SIMILARITY).
FT DOMAIN 5623 5642 243 (BY SIMILARITY).
FT TRANSMEM 5643 5662 244 (BY SIMILARITY).
FT DOMAIN 5663 5682 245 (BY SIMILARITY).
FT TRANSMEM 5683 5702 246 (BY SIMILARITY).
FT DOMAIN 5703 5722 247 (BY SIMILARITY).
FT TRANSMEM 5723 5742 248 (BY SIMILARITY).
FT DOMAIN 5743 5762 249 (BY SIMILARITY).
FT TRANSMEM 5763 5782 250 (BY SIMILARITY).
FT DOMAIN 5783 5802 251 (BY SIMILARITY).
FT TRANSMEM 5803 5822 252 (BY SIMILARITY).
FT DOMAIN 5823 5842 253 (BY SIMILARITY).
FT TRANSMEM 5843 5862 254 (BY SIMILARITY).
FT DOMAIN 5863 5882 255 (BY SIMILARITY).
FT TRANSMEM 5883 5902 256 (BY SIMILARITY).
FT DOMAIN 5903 5922 257 (BY SIMILARITY).
FT TRANSMEM 5923 5942 258 (BY SIMILARITY).
FT DOMAIN 5943 5962 259 (BY SIMILARITY).
FT TRANSMEM 5963 5982 260 (BY SIMILARITY).
FT DOMAIN 5983 6002 261 (BY SIMILARITY).
FT TRANSMEM 6003 6022 262 (BY SIMILARITY).
FT DOMAIN 6023 6042 263 (BY SIMILARITY).
FT TRANSMEM 6043 6062 264 (BY SIMILARITY).
FT DOMAIN 6063 6082 265 (BY SIMILARITY).
FT TRANSMEM 6083 6102 266 (BY SIMILARITY).
FT DOMAIN 6103 6122 267 (BY SIMILARITY).
FT TRANSMEM 6123 6142 268 (BY SIMILARITY).
FT DOMAIN 6143 6162 269 (BY SIMILARITY).
FT TRANSMEM 6163 6182 270 (BY SIMILARITY).
FT DOMAIN 6183 6202 271 (BY SIMILARITY).
FT TRANSMEM 6203 6222 272 (BY SIMILARITY).
FT DOMAIN 6223 6242 273 (BY SIMILARITY).
FT TRANSMEM 6243 6262 274 (BY SIMILARITY).
FT DOMAIN 6263 6282 275 (BY SIMILARITY).
FT TRANSMEM 6283 6302 276 (BY SIMILARITY).
FT DOMAIN 6303 6322 277 (BY SIMILARITY).
FT TRANSMEM 6323 6342 278 (BY SIMILARITY).
FT DOMAIN 6343 6362 279 (BY SIMILARITY).
FT TRANSMEM 6363 6382 280 (BY SIMILARITY).
FT DOMAIN 6383 6402 281 (BY SIMILARITY).
FT TRANSMEM 6403 6422 282 (BY SIMILARITY).
FT DOMAIN 6423 6442 283 (BY SIMILARITY).
FT TRANSMEM 6443 6462 284 (BY SIMILARITY).
FT DOMAIN 6463 6482 285 (BY SIMILARITY).
FT TRANSMEM 6483 6502 286 (BY SIMILARITY).
FT DOMAIN 6503 6522 287 (BY SIMILARITY).
FT TRANSMEM 6523 6542 288 (BY SIMILARITY).
FT DOMAIN 6543 6562 289 (BY SIMILARITY).
FT TRANSMEM 6563 6582 290 (BY SIMILARITY).
FT DOMAIN 6583 6602 291 (BY SIMILARITY).
FT TRANSMEM 6603 6622 292 (BY SIMILARITY).
FT DOMAIN 6623 6642 293 (BY SIMILARITY).
FT TRANSMEM 6643 6662 294 (BY SIMILARITY).
FT DOMAIN 6663 6682 295 (BY SIMILARITY).
FT TRANSMEM 6683 6702 296 (BY SIMILARITY).
FT DOMAIN 6703 6722 297 (BY SIMILARITY).
FT TRANSMEM 6723 6742 298 (BY SIMILARITY).
FT DOMAIN 6743 6762 299 (BY SIMILARITY).
FT TRANSMEM 6763 6782 300 (BY SIMILARITY).
FT DOMAIN 6783 6802 301 (BY SIMILARITY).
FT TRANSMEM 6803 6822 302 (BY SIMILARITY).
FT DOMAIN 6823 6842 303 (BY SIMILARITY).
FT TRANSMEM 6843 6862 304 (BY SIMILARITY).
FT DOMAIN 6863 6882 305 (BY SIMILARITY).
FT TRANSMEM 6883 6902 306 (BY SIMILARITY).
FT DOMAIN 6903 6922 307 (BY SIMILARITY).
FT TRANSMEM 6923 6942 308 (BY SIMILARITY).
FT DOMAIN 6943 6962 309 (BY SIMILARITY).
FT TRANSMEM 6963 6982 310 (BY SIMILARITY).
FT DOMAIN 6983 7002 311 (BY SIMILARITY).
FT TRANSMEM 7003 7022 312 (BY SIMILARITY).
FT DOMAIN 7023 7042 313 (BY SIMILARITY).
FT TRANSMEM 7043 7062 314 (BY SIMILARITY).
FT DOMAIN 7063 7082 315 (BY SIMILARITY).
FT TRANSMEM 7083 7102 316 (BY SIMILARITY).
FT DOMAIN 7103 7122 317 (BY SIMILARITY).
FT TRANSMEM 7123 7142 318 (BY SIMILARITY).
FT DOMAIN 7143 7162 319 (BY SIMILARITY).
FT TRANSMEM 7163 7182 320 (BY SIMILARITY).
FT DOMAIN 7183 7202 321 (BY SIMILARITY).
FT TRANSMEM 7203 7222 322 (BY SIMILARITY).
FT DOMAIN 7223 7242 323 (BY SIMILARITY).
FT TRANSMEM 7243 7262 324 (BY SIMILARITY).
FT DOMAIN 7263 7282 325 (BY SIMILARITY).
FT TRANSMEM 7283 7302 326 (BY SIMILARITY).
FT DOMAIN 7303 7322 327 (BY SIMILARITY).
FT TRANSMEM 7323 7342 328 (BY SIMILARITY).
FT DOMAIN 7343 7362 329 (BY SIMILARITY).
FT TRANSMEM 7363 7382 330 (BY SIMILARITY).
FT DOMAIN 7383 7402 331 (BY SIMILARITY).
FT TRANSMEM 7403 7422 332 (BY SIMILARITY).
FT DOMAIN 7423 7442 333 (BY SIMILARITY).
FT TRANSMEM 7443 7462 334 (BY SIMILARITY).
FT DOMAIN 7463 7482 335 (BY SIMILARITY).
FT TRANSMEM 7483 7502 336 (BY SIMILARITY).
FT DOMAIN 7503 7522 337 (BY SIMILARITY).
FT TRANSMEM 7523 7542 338 (BY SIMILARITY).
FT DOMAIN 7543 7562 339 (BY SIMILARITY).
FT TRANSMEM 7563 7582 340 (BY SIMILARITY).
FT DOMAIN 7583 7602 341 (BY SIMILARITY).
FT TRANSMEM 7603 7622 342 (BY SIMILARITY).
FT DOMAIN 7623 7642 343 (BY SIMILARITY).
FT TRANSMEM 7643 7662 344 (BY SIMILARITY).
FT DOMAIN 7663 7682 345 (BY SIMILARITY).
FT TRANSMEM 7683 7702 346 (BY SIMILARITY).
FT DOMAIN 7703 7722 347 (BY SIMILARITY).
FT TRANSMEM 7723 7742 348 (BY SIMILARITY).
FT DOMAIN 7743 7762 349 (BY SIMILARITY).
FT TRANSMEM 7763 7782 350 (BY SIMILARITY).
FT DOMAIN 7783 7802 351 (BY SIMILARITY).
FT TRANSMEM 7803 7822 352 (BY SIMILARITY).
FT DOMAIN 7823 7842 353 (BY SIMILARITY).
FT TRANSMEM 7843 7862 354 (BY SIMILARITY).
FT DOMAIN 7863 7882 355 (BY SIMILARITY).
FT TRANSMEM 7883 7902 356 (BY SIMILARITY).
FT DOMAIN 7903 7922 357 (BY SIMILARITY).
FT TRANSMEM 7923 7942 358 (BY SIMILARITY).
FT DOMAIN 7943 7962 359 (BY SIMILARITY).
FT TRANSMEM 7963 7982 360 (BY SIMILARITY).
FT DOMAIN 7983 8002 361 (BY SIMILARITY).
FT TRANSMEM 8003 8022 362 (BY SIMILARITY).
FT DOMAIN 8023 8042 363 (BY SIMILARITY).
FT TRANSMEM 8043 8062 364 (BY SIMILARITY).
FT DOMAIN 8063 8082 365 (BY SIMILARITY).
FT TRANSMEM 8083 8102 366 (BY SIMILARITY).
FT DOMAIN 8103 8122 367 (BY SIMILARITY).
FT TRANSMEM 8123 8142 368 (BY SIMILARITY).
FT DOMAIN 8143 8162 369 (BY SIMILARITY).
FT TRANSMEM 8163 8182 370 (BY SIMILARITY).
FT DOMAIN 8183 8202 371 (BY SIMILARITY).
FT TRANSMEM 8203 8222 372 (BY SIMILARITY).
FT DOMAIN 8223 8242 373 (BY SIMILARITY).
FT TRANSMEM 8243 8262 374 (BY SIMILARITY).
FT DOMAIN 8263 8282 375 (BY SIMILARITY).
FT TRANSMEM 8283 8302 376 (BY SIMILARITY).
FT DOMAIN 8303 8322 377 (BY SIMILARITY).
FT TRANSMEM 8323 8342 378 (BY SIMILARITY).
FT DOMAIN 8343 8362 379 (BY SIMILARITY).
FT TRANSMEM 8363 8382 380 (BY SIMILARITY).
FT DOMAIN 8383 8402 381 (BY SIMILARITY).
FT TRANSMEM 8403 8422 382 (BY SIMILARITY).
FT DOMAIN 8423 8442 383 (BY SIMILARITY).
FT TRANSMEM 8443 8462 384 (BY SIMILARITY).
FT DOMAIN 8463 8482 385 (BY SIMILARITY).
FT TRANSMEM 8483 8502 386 (BY SIMILARITY).
FT DOMAIN 8503 8522 387 (BY SIMILARITY).
FT TRANSMEM 8523 8542 388 (BY SIMILARITY).
FT DOMAIN 8543 8562 389 (BY SIMILARITY).
FT TRANSMEM 8563 8582 390 (BY SIMILARITY).
FT DOMAIN 8583 8602 391 (BY SIMILARITY).
FT TRANSMEM 8603 8622 392 (BY SIMILARITY).
FT DOMAIN 8623 8642 393 (BY SIMILARITY).
FT TRANSMEM 8643 8662 394 (BY SIMILARITY).
FT DOMAIN 8663 8682 395 (BY SIMILARITY).
FT TRANSMEM 8683 8702 396 (BY SIMILARITY).
FT DOMAIN 8703 8722 397 (BY SIMILARITY).
FT TRANSMEM 8723 8742 398 (BY SIMILARITY).
FT DOMAIN 8743 8762 399 (BY SIMILARITY).
FT TRANSMEM 8763 8782 400 (BY SIMILARITY).
FT DOMAIN 8783 8802 401 (BY SIMILARITY).
FT TRANSMEM 8803 8822 402 (BY SIMILARITY).
FT DOMAIN 8823 8842 403 (BY SIMILARITY).
FT TRANSMEM 8843 8862 404 (BY SIMILARITY).
FT DOMAIN 8863 8882 405 (BY SIMILARITY).
FT TRANSMEM 8883 8902 406 (BY SIMILARITY).
FT DOMAIN 8903 8922 407 (BY SIMILARITY).
FT TRANSMEM 8923 8942 408 (BY SIMILARITY).
FT DOMAIN 8943 8962 409 (BY SIMILARITY).
FT TRANSMEM 8963 8982 410 (BY SIMILARITY).
FT DOMAIN 8983 9002 411 (BY SIMILARITY).
FT TRANSMEM 9003 9022 412 (BY SIMILARITY).
FT DOMAIN 9023 9042 413 (BY SIMILARITY).
FT TRANSMEM 9043 9062 414 (BY SIMILARITY).
FT DOMAIN 9063 9082 415 (BY SIMILARITY).
FT TRANSMEM 9083 9102 416 (BY SIMILARITY).
FT DOMAIN 9103 9122 417 (BY SIMILARITY).
FT TRANSMEM 9123 9142 418 (BY SIMILARITY).
FT DOMAIN 9143 9162 419 (BY SIMILARITY).
FT TRANSMEM 9163 9182 420 (BY SIMILARITY).
FT DOMAIN 9183 9202 421 (BY SIMILARITY).
FT TRANSMEM 9203 9222 422 (BY SIMILARITY).
FT DOMAIN 9223 9242 423 (BY SIMILARITY).
FT TRANSMEM 9243 9262 424 (BY SIMILARITY).
FT DOMAIN 9263 9282 425 (BY SIMILARITY).
FT TRANSMEM 9283 9302 426 (BY SIMILARITY).
FT DOMAIN 9303 9322 427 (BY SIMILARITY).
FT TRANSMEM 9323 9342 428 (BY SIMILARITY).
FT DOMAIN 9343 9362 429 (BY SIMILARITY).
FT TRANSMEM 9363 9382 430 (BY SIMILARITY).
FT DOMAIN 9383 9402 431 (BY SIMILARITY).
FT TRANSMEM 9403 9422 432 (BY SIMILARITY).
FT DOMAIN 9423 9442 433 (BY SIMILARITY).
FT TRANSMEM 9443 9462 434 (BY SIMILARITY).
FT DOMAIN 9463 9482 435 (BY SIMILARITY).
FT TRANSMEM 9483 9502 436 (BY SIMILARITY).
FT DOMAIN 9503 9522 437 (BY SIMILARITY).
FT TRANSMEM 9523 9542 438 (BY SIMILARITY).
FT DOMAIN 9543 9562 439 (BY SIMILARITY).
FT TRANSMEM 9563 9582 440 (BY SIMILARITY).
FT DOMAIN 9583 9602 441 (BY SIMILARITY).
FT TRANSMEM 9603 9622 442 (BY SIMILARITY).
FT DOMAIN 9623 9642 443 (BY SIMILARITY).
FT TRANSMEM 9643 9662 444 (BY SIMILARITY).
FT DOMAIN 9663 9682 445 (BY SIMILARITY).
FT TRANSMEM 9683 9702 446 (BY SIMILARITY).
FT DOMAIN 9703 9722 447 (BY SIMILARITY).
FT TRANSMEM 9723 9742 448 (BY SIMILARITY).
FT DOMAIN 9743 9762 449 (BY SIMILARITY).
FT TRANSMEM 9763 9782 450 (BY SIMILARITY).
FT DOMAIN 9783 9802 451 (BY SIMILARITY).
FT TRANSMEM 9803 9822 452 (BY SIMILARITY).
FT DOMAIN 9823 9842 453 (BY SIMILARITY).
FT TRANSMEM 9843 9862 454 (BY SIMILARITY).
FT DOMAIN 9863 9882 455 (BY SIMILARITY).
FT TRANSMEM 9883 9902 456 (BY SIMILARITY).
FT DOMAIN 9903 9922 457 (BY SIMILARITY).
FT TRANSMEM 9923 9942 458 (BY SIMILARITY).
FT DOMAIN 9943 9962 459 (BY SIMILARITY).
FT TRANSMEM 9963 9982 460 (BY SIMILARITY).
FT DOMAIN 9983 10002 461 (BY SIMILARITY).

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=97332675; PubMed=9188486;
RA Astry J.M., Jones L.R.;
RT "Functional Co-expression of the canine cardiac Ca2+ pump and
RT phospholamban in Spodoptera frugiperda (Sf21) cells reveals new
RT insights on ATPase regulation.";
RL J. Biol. Chem. 272:15872-15880(1997).
CC -!- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE HYDROLYSIS
CC OF ATP COUPLED WITH THE TRANSLLOCATION OF CALCIUM FROM THE CYTOSOL
CC TO THE SARCOPLASMIC RETICULUM LUMEN. ISOFORM SERCA2A IS INVOLVED
CC IN THE REGULATION OF THE CONTRACTION/RELAXATION CYCLE (BY
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + Ca(2+) (in) = ADP + phosphate +
CC Ca(2+) (out).
CC -!- ENZYME REGULATION: REVERSIBLY INHIBITED BY PHOSPHOLAMBAN (PLN) AT
CC LOW CALCIUM CONCENTRATIONS. DEPHOSPHORYLATED PLN DECREASES THE
CC APPARENT AFFINITY OF THE ATPASE FOR CALCIUM. THIS INHIBITION IS
CC REGULATED BY THE PHOSPHORYLATION OF PLN.
CC -!- SUBUNIT: ASSOCIATED WITH PHOSPHOLAMBAN (PLN).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. SARCOPLASMIC AND
CC ENDOPASMIC RETICULUM.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; SERCA2A/ATP2A2A (SHOWN HERE) AND
CC SERCA2B/ATP2A2B; ARE PRODUCED BY ALTERNATIVE SPLICING IN A TISSUE-
CC SPECIFIC MANNER (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: ISOFORM SERCA2A IS HIGHLY EXPRESSED IN HEART
CC AND SLOW TWITCH SKELETAL MUSCLE. ISOFORM SERCA2B IS WIDELY
CC EXPRESSED.
CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
CC (E1-E2 ATPASES). SUBFAMILY IIA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: U94345; AAC02263.1; -.
DR HSP; P04191; LEU.
DR InterPro; IPR001757; ATPase_E1-E2.
DR InterPro; IPR004014; Cation ATPase.
DR InterPro; IPR000661; H/K Na/K ATPase.
DR InterPro; IPR001454; Hlgase/Hydriase.
DR Pfam; PF00122; E1-E2 ATPase; 1.
DR Pfam; PF00889; Cation ATPase C; 1.
DR Pfam; PF00690; Cation ATPase N; 1.
DR Pfam; PF00702; Hydrolase; 
```

FT TRANSMEM 964 984 10 (BY SIMILARITY).
 FT DOMAIN 995 997 CYTOPLASMIC (BY SIMILARITY).
 FT DOMAIN 370 400 INTERACTS WITH PHOSPHOLAMBAN 1 (BY SIMILARITY).
 FT DOMAIN 787 807 INTERACTS WITH PHOSPHOLAMBAN 2 (BY SIMILARITY).
 FT MOD_RES 351 351 PHOSPHORYLATION (BY SIMILARITY).
 FT METAL 702 702 MAGNESIUM (BY SIMILARITY).
 FT METAL 706 706 MAGNESIUM (BY SIMILARITY).
 FT CA_BIND 304 304 2 (BY SIMILARITY).
 FT CA_BIND 305 305 2 (BY SIMILARITY).
 FT CA_BIND 307 307 2 (BY SIMILARITY).
 FT CA_BIND 309 309 2 (BY SIMILARITY).
 FT CA_BIND 767 767 1 (BY SIMILARITY).
 FT CA_BIND 770 770 1 (BY SIMILARITY).
 FT CA_BIND 795 795 2 (BY SIMILARITY).
 FT CA_BIND 798 798 1 (BY SIMILARITY).
 FT CA_BIND 799 799 1 AND 2 (BY SIMILARITY).
 FT CA_BIND 907 907 1 (BY SIMILARITY).
 SQ SEQUENCE 997 AA; 109652 MW; 266770F599A9E3D3 CRC64;

Query Match 4.0%; Score 99; DB 1; Length 997;
 Best Local Similarity 25.4%; Pred No. 5.2;
 Matches 45; Conservative 28; Mismatches 60; Indels 44; Gaps 11;

QY 21 PRSP-----GMDFIQHLGVCLVALISVGLLSVAACWFL-----PSIIAAASWIITC 69
 DB 820 PRNKPRLISGWL-FFRYLAIGCYGGAATVG--AAAWFIAADGGPRVSFYQLSHFLQC 875
 QY 70 -----VLLCCSKHARCFILLVFLSCGLREGRNALIAAGTIVILGHVENIFHNF 118
 DB 876 KDDNPDFEGVCAIFESPYPMTMALSVLVTIMCNALNSL-SENOSLLRMPWENIW--- 931
 QY 119 KGLLDGMCNLRKAFSIFHPILKKYIEALOMIYGLATPLSVDFDLVSNQTLAVSL 175
 DB 932 ---LVGSIC---LSNSLHLFLIL--YVEPLIFQI-TPL-----NQTQMLMVLKLSL 974

RESULT 7
 ID ATA2_FELCA STANDARD; PRT; 997 AA.
 AC Q00779;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sarcoplasmic/endoplasmic reticulum calcium ATPase 2 (EC 3.6.3.8)
 DE (Calcium pump 2) (SERCA2) (SR Ca(2+)-ATPase 2) (Calcium-transporting
 DE isoform) (Endoplasmic reticulum type, slow twitch skeletal muscle
 DE ATP2A2).
 GN Felis silvestris catus (Cat).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
 OX NCBI_TaxID=9685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RX MEDLINE=92305061; PubMed=1535224;
 RA Gambel A.M., Gallien T.N., Dantzier-Whitworth T., Bowes J.,
 RA Menick D.R.;
 RT "Sequence of the feline cardiac sarcoplasmic reticulum Ca(2+)-
 RT ATPase.";
 RL Biochim. Biophys. Acta 1131:203-206(1992).
 CC -!- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE HYDROLYSIS
 CC OF ATP COUPLED WITH THE TRANSLLOCATION OF CALCIUM FROM THE CYTOSOL
 CC TO THE SARCOPLASMIC RETICULUM LUMEN. ISOFORM SERCA2A IS INVOLVED
 CC IN THE REGULATION OF THE CONTRACTION/RELAXATION CYCLE (BY
 CC SIMILARITY).
 CC -!-CATALYTIC ACTIVITY: ATP + H(2)O + Ca(2+) (In) = ADP + phosphate +
 CC Ca(2+) (Out).
 CC -!- ENZYME REGULATION: REVERSIBLY INHIBITED BY PHOSPHOLAMBAN (PLN) AT
 CC LOW CALCIUM CONCENTRATIONS. DEPHOSPHORYLATED PLN DECREASES THE
 CC APPARENT AFFINITY OF THE ATPASE FOR CALCIUM. THIS INHIBITION IS

CC CC REGULATED BY THE PHOSPHORYLATION OF PLN (BY SIMILARITY).
 CC CC -!- SUBUNIT: ASSOCIATED WITH PHOSPHOLAMBAN (PLN) (BY SIMILARITY).
 CC CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. SARCOPLASMIC AND
 CC CC ENDOPLASMIC RETICULUM.
 CC CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; SERCA2A/ATP2A2A (SHOWN HERE) AND
 CC CC SERCA2B/ATP2A2B; ARE PRODUCED BY ALTERNATIVE SPLICING IN A TISSUE-
 CC CC SPECIFIC MANNER (BY SIMILARITY).
 CC CC -!- TISSUE SPECIFICITY: ISOFORM SERCA2A IS HIGHLY EXPRESSED IN HEART
 CC CC AND SLOW TWITCH SKELETAL MUSCLE. ISOFORM SERCA2B IS WIDELY
 CC CC EXPRESSED.
 CC CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
 CC CC (E1-E2 ATPASES). SUBFAMILY IIA.
 CC CC -----
 CC CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC CC the European Bioinformatics Institute. There are no restrictions on its
 CC CC use by non-profit institutions as long as its content is in no way
 CC CC modified and this statement is not removed. Usage by and for commercial
 CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC CC or send an email to license@isb-sib.ch).
 CC CC -----
 CC CC EMBL; Z11500; CAA77576.1; --
 CC CC PIR; S23444; S23444.
 CC CC HSSP; P04191; LEUL.
 CC CC InterPro; IPR001757; ATPase_E1-E2.
 CC CC InterPro; IPR004014; Cation ATPase.
 CC CC InterPro; IPR000661; H/K Na/K ATPase.
 CC CC InterPro; IPR001454; H1gnase/Hydrlase.
 CC CC Pfam; PF00122; E1-E2 ATPase; 1.
 CC CC Pfam; PF00689; Cation ATPase C; 1.
 CC CC Pfam; PF00690; Cation ATPase N; 1.
 CC CC Pfam; PF00702; Hydrolase; 1.
 CC CC PRINTS; PR00119; CATATPASE.
 CC CC PROSITE; PS00154; ATPASE_E1_E2; 1.
 CC CC Hydrolase; Calcium transport; Transmembrane; Phosphorylation;
 CC CC ATP-binding; Metal-binding; Magnesium; Calcium-binding;
 CC CC Multigene family; Alternative splicing.
 CC CC DOMAIN 1 48 CYTOPLASMIC (BY SIMILARITY).
 CC CC TRANSMEM 49 69 1 (BY SIMILARITY).
 CC CC DOMAIN 70 89 LUMENAL (BY SIMILARITY).
 CC CC TRANSMEM 90 110 2 (BY SIMILARITY).
 CC CC DOMAIN 111 253 CYTOPLASMIC (BY SIMILARITY).
 CC CC TRANSMEM 254 273 3 (BY SIMILARITY).
 CC CC DOMAIN 274 295 LUMENAL (BY SIMILARITY).
 CC CC TRANSMEM 296 313 4 (BY SIMILARITY).
 CC CC DOMAIN 314 756 CYTOPLASMIC (BY SIMILARITY).
 CC CC TRANSMEM 757 776 5 (BY SIMILARITY).
 CC CC DOMAIN 777 786 LUMENAL (BY SIMILARITY).
 CC CC TRANSMEM 787 807 6 (BY SIMILARITY).
 CC CC DOMAIN 808 827 CYTOPLASMIC (BY SIMILARITY).
 CC CC TRANSMEM 828 850 7 (BY SIMILARITY).
 CC CC DOMAIN 851 896 LUMENAL (BY SIMILARITY).
 CC CC TRANSMEM 897 916 8 (BY SIMILARITY).
 CC CC DOMAIN 917 929 CYTOPLASMIC (BY SIMILARITY).
 CC CC TRANSMEM 930 948 9 (BY SIMILARITY).
 CC CC DOMAIN 949 963 LUMENAL (BY SIMILARITY).
 CC CC TRANSMEM 964 984 10 (BY SIMILARITY).
 CC CC DOMAIN 985 997 CYTOPLASMIC (BY SIMILARITY).
 CC CC TRANSMEM 985 997 INTERACTS WITH PHOSPHOLAMBAN 1 (BY
 CC CC SIMILARITY).
 CC CC DOMAIN 787 807 INTERACTS WITH PHOSPHOLAMBAN 2 (BY
 CC CC SIMILARITY).
 CC CC MOD_RES 351 351 PHOSPHORYLATION (BY SIMILARITY).
 CC CC METAL 702 702 MAGNESIUM (BY SIMILARITY).
 CC CC METAL 706 706 MAGNESIUM (BY SIMILARITY).
 CC CC CA_BIND 304 304 2 (BY SIMILARITY).
 CC CC CA_BIND 305 305 2 (BY SIMILARITY).
 CC CC CA_BIND 307 307 2 (BY SIMILARITY).
 CC CC CA_BIND 309 309 2 (BY SIMILARITY).
 CC CC CA_BIND 767 767 1 (BY SIMILARITY).
 CC CC CA_BIND 770 770 1 (BY SIMILARITY).
 CC CC CA_BIND 795 795 2 (BY SIMILARITY).
 CC CC CA_BIND 798 798 1 (BY SIMILARITY).

FT CA BIND 799 799 1 AND 2 (BY SIMILARITY).
 FT CA BIND 907 907 1 (BY SIMILARITY).
 SQ SEQUENCE 997 AA; 109712 MW; CE18D1A1ADA738F CRC64;
 Query Match 4.0%; Score 99; DB 1; Length 997;
 Best Local Similarity 25.4%; Pred. No. 5.2;
 Matches 45; Conservative 28; Mismatches 60; Indels 44; Gaps 11;
 QY 21 PRSP-----GWMDFQHLGVCVLAISVGLLSVNAWFL-----PSIIAAASWITC 69
 DB 820 PRNPKPLISGLW-LFPRYLAIGCVGAATVG---AAAWFIAADGGPRVSFYQLSHFLQC 875
 QY 70 -----VLCCSHKARCFILLVLSGLRGCRNALIAAGTGVILGHVENIFHNF 118
 DB 876 KDNPPDFGVCAIFESPYPMTWALSVLVITMCNALNSL-GENQSLRPPWENIW---- 931
 QY 119 KGLLDGMCNLRKASFSHFPLKRYIEAIOMIYGLATPLSFVDDIVSNQWOTLAVSL 175
 DB 932 ---LVGSIC---LSMSLHFLIL--YVEPLPLIFQI-TPL-----NLTMWLWLKISL 974
 RESULT 8
 AT2A2 HUMAN STANDARD; PRT; 1042 AA.
 AC P16615; P16614;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sarcoplasmic/endoplasmic reticulum calcium ATPase 2 (EC 3.6.3.8)
 DE (Calcium pump 2) (SERCA2) (SR Ca(2+)-ATPase 2) (Calcium-transporting
 DE ATPase sarcoplasmic reticulum type, slow twitch skeletal muscle
 DE isoform) (Endoplasmic reticulum class 1/2 Ca(2+) ATPase).
 GN ATP2A2 OR ATP2B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Eumalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS SERCA2A AND SERCA2B).
 RC TISSUE=Kidney;
 RX MEDLINE=89008384; PubMed=2844796;
 RA Lytton J., MacLennan D.H.;
 RT "Molecular cloning of cDNAs from human kidney coding for two
 RT alternatively spliced products of the cardiac Ca2+-ATPase gene.";
 RL J. Biol. Chem. 263:15024-15031 (1988).
 RN [2]
 RP VARIANTS DD
 RX MEDLINE=99371767; PubMed=10441323;
 RA Sakuntabhai A., Burge S., Monk S., Hovnanian A.;
 RT "Spectrum of novel ATP2A2 mutations in patients with Darier's
 RT disease.";
 RL Hum. Mol. Genet. 8:1611-1619 (1999).
 RN [3]
 RP VARIANTS DD, AND TISSUE SPECIFICITY.
 RX MEDLINE=99371768; PubMed=10441324;
 RA Ruiz-Perez V.L., Carter S.A., Healy E., Todd C., Rees J.L.,
 RA Steijlen P.M., Carmichael A.J., Lewis H.M., Hohl D., Itin P.,
 RA Vahlquist A., Gobbello T., Mazzanti C., Reggazzini R., Nagy G.,
 RA Munro C.S., Strachan T.;
 RT "ATP2A2 mutations in Darier's disease: variant cutaneous phenotypes
 RT are associated with missense mutations, but neuropsychiatric features
 RT are independent of mutation class.";
 RL Hum. Mol. Genet. 8:1621-1630 (1999).
 RN [4]
 RP VARIANTS DD THR-39; ARG-560 AND LEU-765.
 RX MEDLINE=99371769; PubMed=10441325;
 RA Jacobsen N.J.O., Lyons I., Hoogendoorn B., Burge S., Kwok P.-Y.,
 RA O'Donovan M.C., Craddock N., Owen M.J.;
 RT "ATP2A2 mutations in Darier's disease and their relationship to
 RT neuropsychiatric phenotypes.";
 RL Hum. Mol. Genet. 8:1631-1636 (1999).
 RN [5]
 RP VARIANTS DD GLU-23; LYS-357; PHE-495 AND ARG-749.

RX MEDLINE=99178263; PubMed=10080178;
 RA Sakuntabhai A., Ruiz-Perez V., Carter S., Jacobsen N., Burge S.,
 RA Monk S., Smith M., Munro C.S., O'Donovan M.C., Craddock N.,
 RA Kucherlapati R., Rees J.L., Owen M.J., Lathrop G.M., Monaco A.P.,
 RA Strachan T., Hovnanian A.;
 RT "Mutations in ATP2A2, encoding a Ca2+ pump, cause Darier disease.";
 RL Nat. Genet. 21:271-277 (1999).
 CC -!- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE HYDROLYSIS
 CC OF ATP COUPLED WITH THE TRANSLLOCATION OF CALCIUM FROM THE CYTOSOL
 CC TO THE SARCOPLASMIC RETICULUM LUMEN. ISOFORM SERCA2A IS INVOLVED
 CC IN THE REGULATION OF THE CONTRACTION/RELAXATION CYCLE.
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + Ca(2+) (In) = ADP + phosphate +
 CC Ca(2+) (Out).
 CC -!- ENZYME REGULATION: REVERSIBLY INHIBITED BY PHOSPHOLAMBAN (PLN) AT
 CC LOW CALCIUM CONCENTRATIONS. DEPHOSPHORYLATED PLN DECREASES THE
 CC APPARENT AFFINITY OF THE ATPASE FOR CALCIUM. THIS INHIBITION IS
 CC REGULATED BY THE PHOSPHORYLATION OF PLN (BY SIMILARITY).
 CC -!- SUBUNIT: ASSOCIATED WITH PHOSPHOLAMBAN (PLN) (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. SARCOPLASMIC AND
 CC ENDOPLASMIC RETICULUM.
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; SERCA2A/ATP2A2A/CLASS 1/HK2 AND
 CC SERCA2B/ATP2A2B/CLASS 2-4/HK1 (SHOWN HERE); ARE PRODUCED BY
 CC ALTERNATIVE SPLICING. SERCA2 TRANSCRIPTS DIFFER ONLY IN THEIR
 CC 3'UTR REGION AND ARE EXPRESSED IN A TISSUE-SPECIFIC MANNER.
 CC SERCA2A IS A CARDIAC/SLOW TWITCH, MUSCLE SPECIFIC ISOFORM AND
 CC SERCA2B IS A UBIQUITOUS HOUSEKEEPING ISOFORM. SERCA2A HAS A LOWER
 CC AFFINITY FOR CALCIUM AND A HIGHER CATALYTIC TURNOVER RATE.
 CC -!- TISSUE SPECIFICITY: ISOFORM SERCA2A IS HIGHLY EXPRESSED IN HEART
 CC AND SLOW TWITCH SKELETAL MUSCLE. ISOFORM SERCA2B IS WIDELY
 CC EXPRESSED, IN SMOOTH MUSCLE AND NONMUSCLE TISSUES SUCH AS IN ADULT
 CC SKIN EPIDERMIS.
 CC -!- DISEASE: DEFECTS IN ATP2A2 ARE THE CAUSE OF DARIER'S DISEASE (DD)
 CC (ALSO KNOWN AS DARIER-WHITE DISEASE; DAR). DD IS AN AUTOSOMAL
 CC DOMINANTLY INHERITED SKIN DISORDER CHARACTERIZED BY LOSS OF
 CC ADHESION BETWEEN EPIDERMAL CELLS (ACANTHOLYSIS) AND ABNORMAL
 CC KERATINIZATION. PATIENTS WITH MILD DISEASE MAY HAVE NO MORE THAN A
 CC FEW SCATTERED KERATOTIC PAPULES OR SUBTLE NAIL CHANGES, WHEREAS
 CC THOSE WITH SEVERE DISEASE ARE HANDICAPPED BY WIDESPREAD MALODOROUS
 CC KERATOTIC PLAQUES. IN A FEW FAMILIES, NEUROPSYCHIATRIC
 CC ABNORMALITIES SUCH AS MILD MENTAL RETARDATION, SCHIZOPHRENIA,
 CC BIPOLAR DISORDER AND EPILEPSY HAVE BEEN REPORTED. STRESS, UV,
 CC EXPOSURE, HEAT, SWEAT, FRICTION, AND ORAL CONTRACEPTION EXACERBATE
 CC DISEASE SYMPTOMS. PREVALENCE HAS BEEN ESTIMATED AT 1 IN 50000.
 CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
 CC (E1-E2 ATPASES). SUBFAMILY IIA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M23114; AAA53193.1; -
 CC EMBL; M23116; AAA52757.1; -
 CC EMBL; M23115; AAA53194.1; -
 CC EMBL; M23278; AAA52758.1; -
 CC EMBL; M23116; AAA52758.1; JOINED.
 CC PIR; A31981; A31981.
 CC PIR; B31981; B31981.
 CC HSP; P04191; 1EUL.
 CC HSC-2DPAGE; P16614; HUMAN.
 CC Genew; HGNC:812; ATP2A2.
 CC MIM; 108740; -
 CC MIM; 124200; -
 CC InterPro; IPR001757; ATPase_E1-E2.
 CC InterPro; IPR004014; Cation_ATPase.
 CC InterPro; IPR000661; H/K_Na/K_ATPase.
 CC InterPro; IPR001454; Hlgnaase/hydrilase.
 CC Pfam; PF00122; E1-E2_ATPase; 1.
 CC Pfam; PF00689; Cation_ATPase_C; 1.
 CC Pfam; PF00690; Cation_ATPase_N; 1.

CC *ENDOPLASMIC RETICULUM.
 CC -!- ALTERNATIVE PRODUCTS. 2 ISOFORMS; SERCA2A/ATP2A2A AND
 CC SERCA2B/ATP2A2B (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING
 CC IN A TISSUE-SPECIFIC MANNER.
 CC -!- TISSUE SPECIFICITY: ISOFORM SERCA2A IS HIGHLY EXPRESSED IN HEART
 CC AND SLOW TWITCH SKELETAL MUSCLE. ISOFORM SERCA2B IS WIDELY
 CC EXPRESSED.
 CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
 CC (E1-E2 ATPASES). SUBFAMILY IIA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: J04703; AAA31150.1; -
 CC EMBL: X52496; CAA36737.1; -
 CC EMBL: X02814; CAA26583.1; -
 CC PIR: S10335; PWRBMC.
 CC PIR: A33881; A33881.
 CC PIR: A01076; PWRBSC.
 CC HSP: P04191; 1EUL.
 CC InterPro: IPR001757; ATPase_E1-E2.
 CC InterPro: IPR004014; Cation_ATPase.
 CC InterPro: IPR001454; Hlgnaase/hydrlase.
 CC Pfam: PF00122; E1-E2_ATPase; 1.
 CC Pfam: PF00689; Cation_ATPase_C; 1.
 CC Pfam: PF00690; Cation_ATPase_N; 1.
 CC Pfam: PF00702; Hydrolase; 1.
 CC PRINTS: PR00119; CATATPASE.
 CC PROSITE: PS00154; ATPASE_E1_E2; 1.
 CC Hydrolase; Calcium transport; Transmembrane; Phosphorylation;
 KW ATP-binding; Metal-binding; Magnesium; Calcium-binding;
 KW Multigene family; Alternative splicing.
 FT DOMAIN 1 48
 FT TRANSMEM 49 69
 FT DOMAIN 70 89
 FT TRANSMEM 90 110
 FT DOMAIN 111 253
 FT TRANSMEM 254 273
 FT DOMAIN 274 295
 FT TRANSMEM 296 313
 FT DOMAIN 314 756
 FT TRANSMEM 757 776
 FT DOMAIN 777 786
 FT TRANSMEM 787 807
 FT DOMAIN 808 827
 FT TRANSMEM 828 850
 FT DOMAIN 851 896
 FT TRANSMEM 897 916
 FT DOMAIN 917 929
 FT TRANSMEM 930 948
 FT DOMAIN 949 963
 FT TRANSMEM 964 984
 FT TRANSMEM 985 1042
 FT DOMAIN 370 400
 FT DOMAIN 787 807
 FT MOD_RES 351
 FT METAL 702
 FT METAL 706
 FT CA_BIND 304
 FT CA_BIND 305
 FT CA_BIND 307
 FT CA_BIND 309
 FT CA_BIND 309
 FT CA_BIND 767
 FT CA_BIND 770
 FT CA_BIND 795
 FT CA_BIND 798

FT CA_BIND 799 799 1 AND 2 (BY SIMILARITY).
 FT CA_BIND 907 907 1 (BY SIMILARITY).
 FT VARSPIC 994 1042 GXECVQAPQSCSLWACTGCVSWPFFVLLIIVPLVMVYSTDT.
 FT CONFLICT 578 578 NFDLIMS -> AILE (IN ISOFORM SERCA2A).
 FT SEQUENCE 1042 AA; 114704 MW; 4243836D67431575 CRC64;
 Query Match 4.0%; Score 99; DB 1; Length 1042;
 Best Local Similarity 25.4%; Pred. No. 5.4;
 Matches 45; Conservative 28; Mismatches 60; Indels 44; Gaps 11;
 QY 21 PRSP-----GMDFTQHLGCVCLVALISVGLLSVAACWFL-----PSIIAAAASWITC 69
 DB 820 PRNPKEPLISGWL-FFRYLAIGCYGAATVG---AAAWFIADGGGRVSVFYQLSHFLQC 875
 QY 70 -----VLLCCSHKARCFILLVFLSCGLREGRNALIAAGTGTGIVILGHVENIHPNF 118
 DB 876 KEDNPDFFGVDCATFESPYMTALSVLTICMNAISL-SENQSLRPPHENIW--- 931
 QY 119 KGLDGMTCNLRKSGFSHFPLKYYIAQWYIATGLATPLSVFDDLVSMNQTLAVSL 175
 DB 932 ---LVGSIC---LSMSLHFLIL--YVEPLPLIFI--TPLNV---TQWLMVLKISL 974
 RESULT 10
 ID ATA2 RAT STANDARD; PRT; 1043 AA.
 AC P11507; P11508;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sarcoplasmic/endoplasmic reticulum calcium ATPase 2 (EC 3.6.3.8)
 DE (Calcium pump 2) (SERCA2) (SR Ca(2+)-ATPase 2) (Calcium-transporting
 DE ATPase sarcoplasmic reticulum type, slow twitch skeletal muscle
 DE isoform) (Endoplasmic reticulum Class 1/2 Ca(2+) ATPase).
 GN ATP2A2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 [1]
 RN SEQUENCE FROM N.A. (ISOFORMS SERCA2A AND SERCA2B).
 RP TISSUE=Brain;
 RC MEDLINE=89008385; PubMed=2844797;
 RX Guteski-Hamblin A.-M., Greeb J., Shull G.E.;
 RA "A novel Ca2+ pump expressed in brain, kidney, and stomach is encoded
 RT by an alternative transcript of the slow-twitch muscle sarcoplasmic
 RT reticulum Ca-ATPase gene. Identification of cDNAs encoding Ca2+ and
 RT other cation-transporting ATPases using an oligonucleotide probe
 RT derived from the ATP-binding site.";
 RL J. Biol. Chem. 263:15032-15040(1988).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM SERCA2A).
 RC TISSUE=Heart;
 RX MEDLINE=89252068; PubMed=2542094;
 RA Lompre A.M., de la Bastie D., Boheler K.R., Schwartz K.;
 RT "Characterization and expression of the rat heart sarcoplasmic
 RT reticulum Ca2+-ATPase mRNA.";
 RL FEBS Lett. 249:35-41(1989).
 CC -!- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE HYDROLYSIS
 CC OF ATP COUPLED WITH THE TRANSLOCATION OF CALCIUM FROM THE CYTOSOL
 CC TO THE SARCOPLASMIC RETICULUM LUMEN. ISOFORM SERCA2A IS INVOLVED
 CC IN THE REGULATION OF THE CONTRACTION/RELAXATION CYCLE (BY
 CC SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + Ca(2+) (In) = ADP + phosphate +
 CC Ca(2+) (Out).
 CC -!- ENZYME REGULATION: REVERSIBLY INHIBITED BY PHOSPHOLAMBAN (PLN) AT
 CC LOW CALCIUM CONCENTRATIONS. DEPHOSPHORYLATED PLN DECREASES THE
 CC APPARENT AFFINITY OF THE ATPASE FOR CALCIUM. THIS INHIBITION IS
 CC REGULATED BY THE PHOSPHORYLATION OF PLN (BY SIMILARITY).
 CC -!- SUBUNIT: ASSOCIATED WITH PHOSPHOLAMBAN (PLN) (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. SARCOPLASMIC AND
 CC ENDOPLASMIC RETICULUM.

CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; SERCA2A/ATP2A2A AND
 CC SERCA2B/ATP2A2B (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING
 CC IN A TISSUE-SPECIFIC MANNER (BY SIMILARITY).
 CC -!- TISSUE SPECIFICITY: ISOFORM SERCA2A IS HIGHLY EXPRESSED IN HEART
 CC AND SLOW TWITCH SKELETAL MUSCLE. ISOFORM SERCA2B IS WIDELY
 CC EXPRESSED.
 CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
 CC (E1-E2 ATPASES). SUBFAMILY IIA.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: J04022; AAA40785.1; -
 CC EMBL: J04024; AAA40787.1; -
 CC EMBL: J04023; AAA40786.1; -
 CC EMBL: X15635; CAA33645.1; -
 CC PIR: A31982; A31982.
 CC PIR: B31982; B31982.
 CC PIR: S04269; S04269.
 CC HSSP: P04131; LEUL.
 CC InterPro: IPR001757; ATPase_E1-E2.
 CC InterPro: IPR004014; Cation ATPase.
 CC InterPro: IPR000661; H/K Na/K ATPase.
 CC InterPro: IPR001454; HlgNase/hydrase.
 CC Pfam: PF00122; E1-E2_ATPase; 1.
 CC Pfam: PF00689; Cation_ATPase_C; 1.
 CC Pfam: PF00690; Cation_ATPase_N; 1.
 CC Pfam: PF00702; Hydrolase; 1.
 CC PRINTS: PR00119; CATATPASE.
 CC PROSITE: PS00154; ATPASE_E1-E2; 1.
 CC Hydrolase; Calcium transport; Transmembrane; Phosphorylation;
 CC ATP-binding; Metal-binding; Magnesium; Calcium-binding;
 CC Multigene family; Alternative splicing.
 CC DOMAIN 1 48 CYTOPLASMIC (BY SIMILARITY).
 CC TRANSMEM 49 69
 CC DOMAIN 70 89 LUMENAL (BY SIMILARITY).
 CC TRANSMEM 90 110
 CC TRANSMEM 111 253
 CC TRANSMEM 254 273
 CC TRANSMEM 274 295
 CC TRANSMEM 296 313
 CC TRANSMEM 314 756
 CC TRANSMEM 757 776
 CC TRANSMEM 777 786
 CC TRANSMEM 787 807
 CC TRANSMEM 808 827
 CC TRANSMEM 828 850
 CC TRANSMEM 851 896
 CC TRANSMEM 897 916
 CC TRANSMEM 917 929
 CC TRANSMEM 930 948
 CC TRANSMEM 949 963
 CC TRANSMEM 964 984
 CC TRANSMEM 985 1043
 CC TRANSMEM 370 400
 CC TRANSMEM 787 807
 CC TRANSMEM 351 351
 CC TRANSMEM 702 702
 CC TRANSMEM 706 706
 CC TRANSMEM 304 304
 CC TRANSMEM 305 305
 CC TRANSMEM 307 307
 CC TRANSMEM 309 309
 CC TRANSMEM 767 767
 CC TRANSMEM 770 770
 CC TRANSMEM 795 795

FT CA_BIND 798 798 1 (BY SIMILARITY).
 FT CA_BIND 799 799 1 AND 2 (BY SIMILARITY).
 FT CA_BIND 907 907 1 (BY SIMILARITY).
 FT VARSPLIC 994 1043 GRECAQATPKSCSLACTDGIWSPFVLLIMPLVWVYSTD
 TTNSDMFWS -> AILE (IN ISOFORM SERCA2A).
 FT CONFLICT 272 272 W -> T (IN REF. 2).
 FT CONFLICT 288 288 W -> T (IN REF. 2).
 FT CONFLICT 557 557 T -> Q (IN REF. 2).
 SQ SEQUENCE 1043 AA; 114767 MW; 480B476BFD97F390 CRC64;
 Query Match 4.0%; Score 99; DB 1; Length 1043;
 Best Local Similarity 25.4%; Pred. No. 5.4;
 Matches 45; Conservative 28; Mismatches 60; Indels 44; Gaps 11;
 QY 21 PRSP-----GWMDFIQHLGVCCCLVALISVGLSVAAACWFL-----PSIIAAASWITTC 69
 DB 820 PRNPKPLISGWL-PPRYLAIGCYVGAATVG-----AAAWFIAADGGPRVSFYQLSHFLQC 875
 QY 70 -----VLLCCSKHARCFILVFLSCGLUREGNALIAAGTGIVILGHVENIFHNP 118
 DB 876 KEDNPDPEGVDCAFESPYMTWALSVLVTIENCNALNSL-SENQSLRRPPWENIW--- 931
 QY 119 KGLLDGWTNLRKSPSIHFFLLKVKYIEAIQWLYGLATPLSVFDDLVSNQTLAVSL 175
 DB 932 ---LVGSIC-----LSMSLHFLIL--YVEPLIFQI-TPL-----NLQWLMLVKISL 974
 RESULT 11
 AT2A_MOUSE STANDARD; PRT; 1044 AA.
 ID AT2A_MOUSE STANDARD; PRT; 1044 AA.
 AC 055143; Q9R2A9; Q9WUT5;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sarcoplasmic/endoplasmic reticulum calcium ATPase 2 (EC 3.6.3.8)
 DE (Calcium pump 2) (SERCA2) (SR Ca(2+)-ATPase 2) (Calcium-transporting
 DE ATPase sarcoplasmic reticulum type, slow twitch skeletal muscle
 DE isoform) (Endoplasmic reticulum class 1/2 Ca(2+) ATPase).
 GN ATP2A2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC STRAIN=129/SvJ.
 RC MEDLINE=20122169; PubMed=10656932;
 RA Ver Heyen M., Reed T.D., Blough R.E., Zilberman A.L., Loukianov E.,
 RA Van Baelen K., Raeymaekers L., Periasamy M., Wutack F.;
 RT "Structure and organization of the mouse Atp2a2 gene encoding the
 RT sarco(endo)plasmic reticulum Ca(2+)-ATPase 2 (SERCA2) isoforms.";
 RL Mamm. Genome 11:159-163(2000).
 CC -!- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE HYDROLYSIS
 CC OF ATP COUPLED WITH THE TRANSLOCATION OF CALCIUM FROM THE CYTOSOL
 CC TO THE SARCOPLASMIC RETICULUM LUMEN. ISOFORM SERCA2A IS INVOLVED
 CC IN THE REGULATION OF THE CONTRACTION/RELAXATION CYCLE (BY
 CC SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + Ca(2+) (In) = ADP + phosphate +
 CC Ca(2+) (Out).
 CC -!- ENZYME REGULATION: REVERSIBLY INHIBITED BY PHOSPHOLAMBAN (PLN) AT
 CC LOW CALCIUM CONCENTRATIONS. DEPHOSPHORYLATED PLN DECREASES THE
 CC APPARENT AFFINITY OF THE ATPASE FOR CALCIUM. THIS INHIBITION IS
 CC REGULATED BY THE PHOSPHORYLATION OF PLN (BY SIMILARITY).
 CC -!- SUBUNIT: ASSOCIATED WITH PHOSPHOLAMBAN (PLN) (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. SARCOPLASMIC AND
 CC ENDOPLASMIC RETICULUM.
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; SERCA2A/ATP2A2A AND
 CC SERCA2B/ATP2A2B (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING
 CC IN A TISSUE-SPECIFIC MANNER.
 CC -!- TISSUE SPECIFICITY: ISOFORM SERCA2A IS HIGHLY EXPRESSED IN HEART
 CC AND SLOW TWITCH SKELETAL MUSCLE. ISOFORM SERCA2B IS WIDELY
 CC EXPRESSED.
 CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY

CC CC (E1-E2 ATPASES). SUBFAMILY IIA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC EMBL; AJ131821; CAB72436.1; -

CC EMBL; AJ223584; CAAL1450.1; -

CC EMBL; AF029982; AAD01889.1; -

CC EMBL; AJ131870; CAB41017.1; -

CC EMBL; AJ131870; CAB41018.1; -

CC HSP; P04191; 1EUL.

CC MGD; MGI:88110; Atp2a2.

CC InterPro; IPR001757; ATPase E1-E2.

CC InterPro; IPR004014; Cation_ATPase.

CC InterPro; IPR001454; Hignase/hydrilase.

CC Pfam; PF00122; E1-E2_ATPase; 1.

CC Pfam; PF00689; Cation_ATPase_C; 1.

CC Pfam; PF00690; Cation_ATPase_N; 1.

CC Pfam; PF00702; Hydrolase; 1.

CC PRINTS; PR00119; CATATPASE.

CC PROSITE; PS00154; ATPASE E1 E2; 1.

CC Hydrolase; Calcium transport; Transmembrane; Phosphorylation;

CC ATP-binding; Metal-binding; Magnesium; Calcium-binding;

CC Multigene family; Alternative splicing.

CC DOMAIN 1 48

CC TRANSMEM 49 69

CC DOMAIN 70 89

CC TRANSMEM 90 110

CC DOMAIN 111 253

CC TRANSMEM 254 273

CC DOMAIN 274 295

CC TRANSMEM 296 313

CC DOMAIN 314 756

CC TRANSMEM 757 776

CC DOMAIN 777 786

CC TRANSMEM 787 807

CC DOMAIN 808 827

CC TRANSMEM 828 850

CC DOMAIN 851 896

CC TRANSMEM 897 916

CC DOMAIN 917 929

CC TRANSMEM 930 948

CC DOMAIN 949 963

CC TRANSMEM 964 984

CC DOMAIN 985 1044

CC TRANSMEM 370 400

CC DOMAIN 787 807

CC INTERACTS WITH PHOSPHOLAMBAN 2 (BY

CC SIMILARITY).

CC MOD_RES 351

CC METAL 702

CC METAL 706

CC CA_BIND 304

CC CA_BIND 305

CC CA_BIND 307

CC CA_BIND 309

CC CA_BIND 767

CC CA_BIND 770

CC CA_BIND 795

CC CA_BIND 798

CC CA_BIND 799

CC CA_BIND 907

CC CA_BIND 907

CC VARSPIC 995 1044

CC GKECVQPKSKSCSIASCTDGIWSPEVLLIMPLVWVYSTD

CC TNFSDMFWS --> AILE (IN ISOFORM SERCA2A).

CC TNFSDMFWS; 06AV53982116C421 CRC64;

CC SEQUENCE 1044 AA; 114857 MW; 114857 MW; 925F4BD18C473C5B CRC64;

CC Query Match 4.0%; Score 99; DB 1; Length 1044;

CC Best Local Similarity 25.4%; Pred. No. 5.4;

Matches 45; Conservative 28; Mismatches 60; Indels 44; Gaps 11;

QY 21 PRSP-----GWMDFQHLGVCLVALISVGLLSVAACWFL-----PSIIAAASWIITC 69

DB 820 PRNPREPLISGWL-PPRYLAIGCYGAATVG---AAAWFIAADGGGRVSFYQLSHFLQC 875

QY 70 -----VLLCCSKHARCFILLVFLSCGLREGNRLIAAGTGIVILGHVENIHFNF 118

DB 876 KEDNPDFDGVCAIFESPYPMTWALSVLVTIENCNALNSL-SENQSLRLPPWENIW--- 931

QY 119 KGLLDGMTCNLRKAKSFHFPFLKVKYIEAIQWYIGLATPLSVFDDLVSMNQTLAVSL 175

DB 932 ---LVGSIC-----LSMSLHFLIL--YVEPLPLIFQI-TPL-----NLQWLMVLKISL 974

RESULT 12

Y069 CHLTR

ID Y069 CHLTR STANDARD; PRT; 451 AA.

AC 084072;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Probable metal transport system membrane protein CT069.

GN CT069.

OS Chlamydia trachomatis.

OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.

OX NCBI_TaxID=813;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=D/UN-3/Cx;

RA MEDLINE=99000809; PubMed=9784136;

RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,

RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,

RA Davis R.W.;

RT "Genome sequence of an obligate intracellular pathogen of humans:

RT Chlamydia trachomatis."

RL Science 282:754-759(1998).

CC -!- FUNCTION: PART OF AN ATP-DRIVEN TRANSPORT SYSTEM

CC CT067/CT068/CT069/CT070 FOR A METAL.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane

CC (Probable).

CC -!- SIMILARITY: BELONGS TO THE ABC-3 SUBFAMILY OF INTEGRAL MEMBRANE

CC PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC EMBL; AE001281; AAC67660.1; -

CC InterPro; IPR001626; Abctransprtr3.

CC InterPro; IPR001367; HTH_DtXR.

CC Pfam; PF00950; ABC-3; 1.

CC Pfam; PF02742; Fe_dep_repr_C; 1.

CC SMART; SM00529; HTH_DtXR; 1.

CC KW Hypothetical protein; Transmembrane; Transmembrane; Inner membrane;

CC Complete proteome.

CC TRANSMEM 14 34 POTENTIAL.

CC TRANSMEM 38 58 POTENTIAL.

CC TRANSMEM 70 90 POTENTIAL.

CC TRANSMEM 100 120 POTENTIAL.

CC TRANSMEM 145 165 POTENTIAL.

CC TRANSMEM 192 212 POTENTIAL.

CC TRANSMEM 233 253 POTENTIAL.

CC TRANSMEM 259 289 POTENTIAL.

CC SEQUENCE 451 AA; 51231 MW; 925F4BD18C473C5B CRC64;

Query Match 3.9%; Score 96; DB 1; Length 451;

Best Local Similarity 18.8%; Pred. No. 3.6;

Matches 74; Conservative 54; Mismatches 112; Indels 154; Gaps 18;

Search completed: June 21, 2003, 13:24:10
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2003, 13:19:27 ; Search time 93 Seconds
(without alignments)
1166.773 Million cell updates/sec

Title: US-09-713-098-2

Perfect score: 2448

Sequence: 1 MGITWSTGDIPLSLWEIVVS.....LPVLKMRKKQMDWASADKS 470

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2448	100.0	470	4	Q9H295
2	1438.5	58.8	367	11	Q9D619
3	233.5	9.5	639	4	Q96M03
4	207	8.5	706	4	Q96M70
5	189.5	7.7	551	4	Q9BR26
6	168.5	6.9	498	11	Q9D611
7	158.5	6.5	684	5	Q9VU52
8	135.5	5.5	334	5	Q961F6
9	122.5	5.0	689	16	Q9KY46
10	113.5	4.6	351	5	Q17520
11	113	4.6	315	10	Q8W4R6
12	104.5	4.3	496	2	Q93CP7
13	104	4.2	320	2	Q9RNE1
14	104	4.2	401	10	Q9FLR8
15	104	4.2	1971	4	Q9NTT5
16	103	4.2	638	5	Q917H4

17	103	4.2	569	5	Q8SRG1
18	102.5	4.2	407	16	Q9ZCV6
19	102	4.2	391	16	Q9PLZ4
20	102	4.2	536	2	O06762
21	102	4.2	741	16	Q8RFP8
22	101	4.1	635	2	Q9F443
23	101	4.1	645	5	Q21688
24	100	4.1	505	16	Q98I66
25	99.5	4.1	470	16	Q9JSS5
26	99.5	4.1	690	10	Q9SJA4
27	99.5	4.1	1444	5	Q95XW0
28	99	4.0	510	11	Q63080
29	98.5	4.0	578	5	Q9VCG4
30	98	4.0	334	5	Q9NSD1
31	98	4.0	573	16	Q8YRB4
32	98	4.0	598	2	Q50866
33	97.5	4.0	1065	16	O31501
34	97	4.0	592	16	Q97IH2
35	97	4.0	614	16	O31615
36	97	4.0	626	16	O51574
37	97	4.0	634	5	Q8TOL2
38	96.5	3.9	732	5	O18050
39	96.5	3.9	787	10	Q8VXB7
40	96	3.9	405	16	O24865
41	96	3.9	700	5	Q9UAB9
42	96	3.9	1465	3	Q9P5N0
43	95.5	3.9	470	16	Q926R4
44	95.5	3.9	658	5	Q9U3N1
45	95	3.9	346	8	Q954E2

ALIGNMENTS

RESULT 1

Q9H295 PRELIMINARY; PRT; 470 AA.
 ID Q9H295;
 AC Q9H295;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE DC-specific transmembrane protein (Putative IL-4 induced protein DE FIND)
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20586129; PubMed=11169400;
 RA Hartgers F.C., Vissers J.L.M., Looman M.W.G., van Zoelen C., RA Huffin C., Fidor C.G., Adema G.J.;
 RT "DC-STAMP, a novel multimeric membrane-spanning molecule preferentially expressed by dendritic cells";
 RL Eur. J. Immunol. 30:3585-3590(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21243722; PubMed=11345586;
 RA Staeger H., Brauchlin A., Schoedon G., Schaffner A.;
 RT "Two novel genes FIND and LIND differentially expressed in deactivated RT and Listeria-infected human macrophages";
 RL Immunogenetics 53:105-113(2001).
 DR EMBL; AF305068; AAC39167.1; -;
 DR EMBL; AF277290; AAL02152.1; -;
 DR InterPro; IPR001211; PhospholipaseA2.
 DR InterPro; IPR001680; WD40.
 DR PROSITE; PS00118; PA2_HIS; UNKNOWN 1.
 DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
 KW Transmembrane.
 SQ SEQUENCE 470 AA; 53392 MW; EA2B858FD2C7560C CRC64;
 Query Match 100.0%; Score 2448; DB 4; Length 470;
 Best Local Similarity 100.0%; Pred. No. 1.3e-205;

Q8SRG1	encephalito
Q9ZCV6	ricketeia
Q9PLZ4	campylobact
O06762	mycoplasma
Q8RFP8	fusobacteri
Q9F443	streptococc
Q21688	caenorhabdi
Q98I66	rhizobium 1
Q9JSS5	chlamydia p
Q9SJA4	arabidopsis
Q95XW0	caenorhabdi
Q63080	rattus norv
Q9VCG4	drosophila
Q9NSD1	caenorhabdi
Q8YRB4	anabaena sp
Q50866	myxococcus
O31501	bacillus su
Q97IH2	clostridium
O31615	bacillus su
O51574	borrelia bu
Q8TOL2	drosophila
O18050	caenorhabdi
Q8VXB7	oryza sativ
O24865	heicobacte
Q9UAB9	leishmania
Q9P5N0	schizosacch
Q926R4	listeria in
Q9U3N1	caenorhabdi
Q954E2	petauroides

Matches 470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGWTSCTDIFLSLWEIYVSPSPGMDPFIQHLGVCCCLVALISVGLLSVAACWFLPSIIA 60
 DB 1 MGWTSCTDIFLSLWEIYVSPSPGMDPFIQHLGVCCCLVALISVGLLSVAACWFLPSIIA 60
 QY 61 AAASWITTCVLLCCSKHARCFILLVFLSCGLREGNRLIAAGTGIVILGHVENIFHNPKG 120
 DB 61 AAASWITTCVLLCCSKHARCFILLVFLSCGLREGNRLIAAGTGIVILGHVENIFHNPKG 120
 QY 121 LLDGMCNLRKAKSFHFPFLKKYIEAIOIYGLATPLSVFDDLVSNQTLAVSLFSPSH 180
 DB 121 LLDGMCNLRKAKSFHFPFLKKYIEAIOIYGLATPLSVFDDLVSNQTLAVSLFSPSH 180
 QY 181 VLEAQLNSKGEVSLVLYOMATTTVEVSLSGOKLLAFAGLSVLVLTGLFMKRFGLPCGW 240
 DB 181 VLEAQLNSKGEVSLVLYOMATTTVEVSLSGOKLLAFAGLSVLVLTGLFMKRFGLPCGW 240
 QY 241 KYENIYITRFQVOPDERERHQRPCVLPNKEERKYYIIPFWPTPKERKNLGLFLPI 300
 DB 241 KYENIYITRFQVOPDERERHQRPCVLPNKEERKYYIIPFWPTPKERKNLGLFLPI 300
 QY 301 LIHLICWVLFPAADVLLYRLIFSVSKQFOSLPGEVHLKLGKQGTODIHDSSFNISV 360
 DB 301 LIHLICWVLFPAADVLLYRLIFSVSKQFOSLPGEVHLKLGKQGTODIHDSSFNISV 360
 QY 361 FEPNCIPKPKFLLSETWVPLSVILLVLMGLSSILMQLKILVSASFYPSVERKRIQYL 420
 DB 361 FEPNCIPKPKFLLSETWVPLSVILLVLMGLSSILMQLKILVSASFYPSVERKRIQYL 420
 QY 421 HAKLLKRSKQPLGVEKRLSLYLTKIHFWLPVLMKIRKQMDMASADKS 470
 DB 421 HAKLLKRSKQPLGVEKRLSLYLTKIHFWLPVLMKIRKQMDMASADKS 470

RESULT 2

Q9D619 ID Q9D619 PRELIMINARY; PRT; 367 AA.

AC Q9D619

DT 01-JUN-2001 (TReMBLrel. 17, Created)

DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)

DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)

DE 4833414107RIK.

GN 4833414107RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=HEAD;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaide C., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,

RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.;"

RL Nature 409:685-690(2001).

DR EMBL; AW014697; BAB29508.1; -.

DR MGD:1923016; 4833414107RIK.

SQ SEQUENCE 367 AA; 42104 MW; 964ADCF585B543B1 CRC64;

Query Match 58.8%; Score 1438.5; DB 11; Length 367;

Best Local Similarity 58.3%; Pred. No. 1.6e-117;

Matches 273; Conservative 39; Mismatches 53; Indels 103; Gaps 1;

QY 1 MGWTSCTDIFLSLWEIYVSPSPGMDPFIQHLGVCCCLVALISVGLLSVAACWFLPSIIA 60
 DB 1 MELWTGLTISFLRLMGTYVFRSPSWLDFIQHLGVCCVFAFLSVLSFAFYILPPVAL 60
 QY 61 AAASWITTCVLLCCSKHARCFILLVFLSCGLREGNRLIAAGTGIVILGHVENIFHNPKG 120
 DB 61 LSSVWMITTCVLLCCSKARARCFILLVFLSCGLREGNRLIAAGTGIVIFGVENIFYNFRG 120
 QY 121 LLDGMCNLRKAKSFHFPFLKKYIEAIOIYGLATPLSVFDDLVSNQTLAVSLFSPSH 180
 DB 121 LLDGMCNLRKAKSFHFPFLKKYIEAIOIYGLATPLSVFDDLVSNQTLAVSLFSPSH 180
 QY 181 VLEAQLNSKGEVSLVLYOMATTTVEVSLSGOKLLAFAGLSVLVLTGLFMKRFGLPCGW 240
 DB 181 ALEAHMNDTRGEVLGVHLHNVVTTTELLTSVGOKLLALAGLLLVSTGLFKFLGPGCW 240
 QY 241 KYENIYITRFQVOPDERERHQRPCVLPNKEERKYYIIPFWPTPKERKNLGLFLPI 300
 DB 241 KYENIYITRFQVOPDERERHQRPCVLPNKEERKYYIIPFWPTPKERKNLGLFLPI 300
 QY 301 LIHLICWVLFPAADVLLYRLIFSVSKQFOSLPGEVHLKLGKQGTODIHDSSFNISV 360
 DB 301 LTYLYWVLFPAADVLLYRLISMMKQFOSLPGLGVHLKLGKGE----- 343
 QY 361 FEPNCIPKPKFLLSETWVPLSVILLVLMGLSSILMQLKILVSASFYPSVERKRIQYL 420
 DB 344 ----- 343
 QY 421 HAKLLKRSKQPLGVEKRLSLYLTKIHFWLPVLMKIRKQMDMASAD 468
 DB 344 -----IHLWFPVLMKIRKQMDMASAD 365

RESULT 3

Q96M03 ID Q96M03 PRELIMINARY; PRT; 639 AA.

AC Q96M03

DT 01-DEC-2001 (TReMBLrel. 19, Created)

DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)

DE CDNA FLJ32934 fis, clone TEst12007480.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RA Ishibashi T., Kanehori K., Yoshida M., Watanabe S., Ishida S., Ono Y.,

RA Hotuta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J.,

RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,

RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,

RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,

RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,

RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,

RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;

RT "NEO human cDNA sequencing project.;"

RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK057496; BAB71511.1; -.

SQ SEQUENCE 639 AA; 71728 MW; E68937CDB0F941D5 CRC64;

Query Match 9.5%; Score 233.5; DB 4; Length 639;

Best Local Similarity 20.8%; Pred. No. 6.2e-12;

Matches 127; Conservative 88; Mismatches 194; Indels 201; Gaps 24;

QY 1 MGWTSCTDIFLSLWEIYVSPSPGMDPFIQHLGVCCCLVALISVGLLSVAACWFLPSI 58

```

Db 27 VGGFTLGLSLATAYAGLELLVEGHP-WGCLV---GTLTLAFLSLG----- 69
Qy 59 TAAASWITTCVLLCCSHARCFILLVFLSCGLREGNIALIAAGTGVILGHVENIFHNF 118
Db 70 -----MGFSQRVATVLLLPQAFSRQGTLLVAAFLVGLQGPCANTLRF 116
Qy 119 KLLDGMTCNL-----RAKSPSIHPL-----LKYIE 146
Db 117 TRASEAVACGAELALNOTAEVLQRAQ-----PLVSALNKIKAIARKTEVADRVRKFFR 171
Qy 147 AI-----OWIY----- 152
Db 172 SIMDGVKHARLRNWNWLLHIGVNCSELNPVLCARVFDADKOSCMWVPOAYHLC 231
Qy 153 -----GLATPLSVFDDLVSW-----NOTLAVSL-----FSPSHVL 182
Db 232 VYLMPPFKLALCGLASLVQFCVPIPKYIQPFLQTGTGTPVIQLNVRQEFENMTATHHF 291
Qy 183 EAQNDKSGEVLVSVLYQWATTEVLSLQKUL-----LAFAGLSVLVLTGLFMKR-PL 235
Db 292 SVDLNASRS-----LSQVA---MDLHEAVSMKLRVREALALMGFTTPLLVLVLLYQALFY 344
Qy 236 GPC--GW-KYENIYITROFVQDERERHQORCVLPINKEERKYYIIPFTWPTPKRXN 292
Db 345 RYCYLNWDHNYIYITRFLRMEAVRSTAGLPTVLPLSAHEARRYIPPGSIFLSQWEKFF 404
Qy 293 LGLFPLILHLCIWLFAADVLLYRLIFSVSKQFS--LPGFEVHLKHGKQG--TOD 349
Db 405 YILEFTNLRLHLLVFLVFLDYAVFWVLDLARHQLQGEIVARSPLVSLTVEGTGYAGN 464
Qy 350 IHDD--SSF-----NISVFEPNCIPKPLSETWVPLSVILLILVLMGLSSILMLQK 401
Db 465 IYRDLVSAFVQLQGNISILSRCLLRSEPDPSTGVIVGVYGLCFITLPGSVYSLR 524
Qy 402 ILVSAFSPVSRKRIQVHLAKLKKRSKQPIG-----EVKRL-----SLYTKI----- 447
Db 525 RVICASYPSREQEERISLYNVLRLSRTNLLAALHRSVRRRAADQHRSAFVLASRCP 584
Qy 448 -----HFWL 451
Db 585 LGPFVSHFWL 594

RESULT 4
Q56M70 PRELIMINARY; PRT; 706 AA.
AC Q56M70;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE CDNA FLJ32785 fis, clone TEST12002251.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA Hotuta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J., Chiba Y.,
RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
RA Sugiyama T., Irie R., Osuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Iseogi T.,
RT "NEDO human cDNA sequencing Project".
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK057347; BABY1440.1; -.
DR InterPro; IPR001841; Znf ring.
SQ SEQUENCE 706 AA; 80677 MW; BAB5BD4599EBC19E CRC64;

Query Match 8.5%; Score 207; DB 4; Length 706;
Best Local Similarity 19.5%; Pred. No. 1.4e-09;
Matches 123; Conservative 86; Mismatches 172; Indels 250; Gaps 23;

Qy 6 SGTDFLSLWEIYVSPSPQWDFIOHLGVCCCLVALISVGLLSVAACWFLPSIIAAAAASW 65
Db 56 AGGLAIGLIFOLLVNP-----MNIYEQIMFLYSLVGLG-----AMGW 94
Qy 66 IITCVLLCCSHARCFILLVFLSCGLREGNIALIAAGTGVILGHVENIFHNFKGLDGM 125
Db 95 -----GTSPIRCASLLLPQMLGKGRFLVGLVLAALAIYVGPVANRLHNLNNVIASL 147
Qy 126 TC-----NLRAKSPSIHPLKKYIEALQNIYGLATPLSVFDDLVSNQTLAV---SL 175
Db 148 GCTVELQINTRA-----AMRISTAPLRAMPFKDLSKELLRAETRNI 190
Qy 176 FSPSHVLEAQLNDSKGEV-----LSV--LYQMATTTEVL 208
Db 191 SATFEDLDAQVSETGYTPEDTWDSETAQGRARQAPASRLHLSLSTOKMYELKTLKRCY 250
Qy 209 SLQCKLLA-----FAGLSLVL----- 224
Db 251 VVQNAILSCRRWFRDKHQCKHIWVPLLTHLLCLPMKFKFCGIAKVMVWCNRIPVE 310
Qy 225 ----- 225
Db 311 GNFGQTYDSINQIRGLDGEFSANIDFKEBKAGVLGLNTSWERVSVEVDYVYVQEARL 370
Qy 226 GTGLPMKRFLGCGW-----KYENIYITROFVQDERERHQORCVL 267
Db 371 ENALGLHLVLLSCTFLVLHASFYSYNDHIFDNIIYSTYFCQIDDRKKLGKRTLL 430
Qy 268 PLNKERRKRYVIPTFWPT--PKERNLGL--LFFLP--LIHLC--IWLFAADVLL 317
Db 431 PLRKAB-EKTVIFPC-KPTIQASENSNVRELLETPIPLLLVVLGCLDHALYISFTIR 488
Qy 318 YRLIFSVSKQFOSLPGFEVHLKHGK-----OGTQDIHDSFSNISVPEPN---CIPK 368
Db 489 HHSFLQYS--FRSSHLEK--KVGSDLLARLARKTKTIGALTSS--ETVMESNNMPCLPQ 542
Qy 369 PKFLLSETWVPLSVILLILVLMGLSSILMLQKILVSAFSPVSRKRIQVHLAKLKKR 428
Db 543 PVGLDARAYWRAAVPIGLLVCLLQAFGYLRSRVIAAFYFPKREKRIILFLYNDLLKKR 602
Qy 429 S-----KQPLGEVKER 439
Db 603 AAFATKLRRRAILRRERQKAPRHPPLADILHR 633

RESULT 5
Q9BR26 PRELIMINARY; PRT; 551 AA.
AC Q9BR26;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE DJ257E24.3 (Novel protein) (Fragment).
GN DJ257E24.3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd D.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL034424; CAC36102.1; -.
DR PRINTS; PR01586; TWIKCHANNEL.
FT NON_TER 1
SQ SEQUENCE 551 AA; 60104 MW; 38A47A2C6B9546AE CRC64;

Query Match 7.7%; Score 189.5; DB 4; Length 551;
Best Local Similarity 20.6%; Pred. No. 3.6e-08;
```



```

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.M.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Butris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003538; AAF49838.1; -.
DR FlyBase; FBgn0036347; CG11281.
DR InterPro; IPR001841; Znf_ring.
DR SMART; SM00184; RING_1.
DR SMART; SM00184; RING_1.
SQ SEQUENCE 684 AA; 79599 MW; 504526517354147C CRC64;

Query Match 6.5%; Score 158.5; DB 5; Length 684;
Best Local Similarity 17.1%; Pred. No. 2.4e-05;
Matches 91; Conservative 83; Mismatches 157; Indels 201; Gaps 17;

QY 66 IITCVLLCCSHARCIFILLVFLSCGLRGRNALIAAGTGIVLGHVENIFHNFKGLDGM 125
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 74 LVTGTIFMLSLFVRAVILLIFVALVKGSGRTYLRVAVAFITSGPIANLVENAGEVARVF 133
QY 126 TC-----NLRKSFSTHFPLLKKYIEAIQWI----- 151
Dy 134 VCTTVLYNLSKTRPDLMAKDFNTLKHMRGVDSEIRHTFVELQVLDLVLYAVENSIE 193
QY 152 ---YG-----LATPLSV----- 160
Dy 194 DEKYGDKNTPIYERWGETSRMNVSEICNGGKELPTPAQVQERQRMNRCKQLRSG 253
QY 161 -----FDPLVS-----WNQTLAV-----SLFSPSHVL----- 182
Dy 254 HRACLEVPNGRYKCTTNFPSPMIKAIICWPRYVDIICELDLFGNPDKIDPSAVVPQNF 313
QY 183 -----EAQINDSKEVLSVLYQMATTEVLSLG----- 211
Dy 314 ETVVELLKAEKKFLDNSQIV-VNVEIKDEQAFKSQLKSABRTGQAFKEDFERQKRIFNK 372
QY 212 -----QKLLAFAGLSVLGLTGLFMKRLPGCGWKYENIYITQVQFDEREHRQPCV 266
Dy 373 VNGILOKILCLFMURWTVTSINYVK-YLNDV--EFDNFYITKPKVQDQRKEORIDAI 429
QY 267 LPLNKEERRKYIITFTWPTPKRKNLGLFPLILHLCIWLFLFAVDYLLYRLFVSVK 326
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 430 LPLRTYKSKYIDVD-----H-----IFSRTHSTTIVRKRSK 463
QY 327 -QFQSLPGFEVHLKHEKQGT---QDIHDSFNI-----SVFPCNICPKPLISET 376
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 464 IVVQDQGEHEVRNIGSVGMARLLRTTMH--NFNIHEKVSTSLSKCECLPNAHVLPKKM 521
QY 377 WYPLSVILLILVWLGLLSLQMLKLVSAFSPYSVERKRIQVHLAKLKKR 428
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 522 YYQLILLYLIIIVLYQSTTFLMRRVICSFYYKREKQRIELFLYNILRNR 573

RESULT 8
Q961F6 PRELIMINARY; PRT; 334 AA.
ID Q961F6; AC Q961F6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE GH26634p.
GN CG6845.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY051619; AAK93043.1; -.
DR FlyBase; FBgn0035099; CG6845.
SQ SEQUENCE 334 AA; 38756 MW; E443D7E15D5A67D0 CRC64;

Query Match 5.5%; Score 135.5; DB 5; Length 334;
Best Local Similarity 22.6%; Pred. No. 0.0011;
Matches 55; Conservative 54; Mismatches 87; Indels 47; Gaps 12;

QY 241 KYENIYITRQFVQFDER-ERHQORPCVLPNKEERKYYIITFTWPTPKRKNLGLF-FL 298
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 5 QFQNVFLTKLSDIDRRHEKHGVD-LLPLHQLERAKYMKLTSL-----RLTLEFV 55
QY 299 PILHLCIW-----LFAA--VDYLYRLIIFSVMKQFSLPGFEVHLKHEKQGT---TQ 348
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 56 SIVENACFMATTCLOLQFAICFLDYGFLWLLATISLHGHOETGLEVPAYVDLEIKGGFVA 115
QY 349 DIIHDSFNIISVFEF-----NCIPKPKLSETWVPLSVILLILVWLGLL---SS 395
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 116 DVNRGIA---NAPRPLTQKSILOVNPCLP---VKPDVABYLIILLCLLAWLILLAE 169
QY 396 ILMQLKILVSASFYSVERKRIQYHLAKLKKRS-----KQPLGEVKRRLSLYLTKIH 448
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 170 YILRTHLIMAHFYPKAKERGMFLYNMISEDTSTFKIRRRKRNEFNTRNVKTRNY 229
QY 449 FWL 451
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 230 TWL 232

RESULT 9
Q9KY46 PRELIMINARY; PRT; 689 AA.
ID Q9KY46; AC Q9KY46;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative integral membrane transport protein.
GN SCO2344 OR SCC8A.02C.
OS Streptomyces coelicolor.

```


Db 180 FOKLPLPSYIYEAPILVLSLTHATISVVYFIVLVLESFVGVLYLFINIVKQMKHK 239

QY 329 QSLPGFEVHLKHGEKQGTQDIIHDSSENFISFENPCPKPFLSETWVPLSVILLIIV 388

Db 240 MSPKTFELQKK-----FIIT-----LLIQVSIPMICFIITLI 271

QY 389 MGLSLSSILMQLKILVSA-----SPYSPVERKIQYLHA-----KLKKESKOPLGSRV 436

Db 272 YIGFAYLINYNOGLNNATLAFSCHGVSSTIALIHALHAPYREYAQDILLKLRMSFVEV 331

QY 437 KR 438

Db 332 SQ 333

RESULT 11

Q8W4R6 PRELIMINARY; PRT; 315 AA.

AC Q8W4R6

DT 01-MAR-2002 (Tremblrel. 20, Created)

DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)

DE AT327760/MGF10 16.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;

RT "Arabidopsis cDNA clones."

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY060578; AAL31204.1; -- 323575A24DC4195 CRC64;

SQ SEQUENCE 315 AA; 36715 MW; 3615 MW; 323575A24DC4195 CRC64;

Query Match 4.6%; Score 113; DB 10; Length 315;

Best Local Similarity 23.4%; Pred. No. 0.092;

Matches 79; Conservative 35; Mismatches 127; Indels 96; Gaps 15;

QY 6 SGTDFL--SLWEIYVSPSPGWMDFIOHLGVCCILVALISVGLLSVAACWFLPSTIAAAA 63

Db 53 NGNDVLCIDDVWRPCFERIHGWLGLGRVILGFCFLA-----NNIARPA 96

QY 64 S-----WIITVLLCCSHKARCFILVFLSC--GLREGNALIAAGTGIVILGH 110

Db 97 NRGWRIYYTQWTFLIAIVGMSG----LLSIYGCLOYKQKNTGIADQVGI----D 148

QY 111 VENIEHNFKGLDGMTCNLRAKSPSIHFPPLKKYIEAIQWYGLATPLSVDFDLVSNQOT 170

Db 149 AENGERS--PLIDGNWVSFEKRTSGSEALKSYVHLFOIYQMGAGAVLTDTSYW--T 204

QY 171 LAVSLFSPSHVLEAQNDSKGEVLSVLYQMATTEVLSLQGLKALFAGLSVLLGTGLF 230

Db 205 VIFPFLS-----LQD-----YEMSPMTNLHT-----SNLVLLLLIDTF 237

QY 231 MKRFLGPCGKYNENIYITROFQDERERHQRPCVLPNKERRKYVIPTFWTPKER 290

Db 238 LNRKFPPLFRFSYFIWTCGFLQ-----WIL-----HMFISVGWYYP--F 277

QY 291 KNIGLFFLPI-----LIHLCIWLFAVDYLLYRLI 321

Db 278 LNLSDMAPVWVLLVALLHLPSYGLFALIVKIKYLI 314

RESULT 12

Q93CP7 PRELIMINARY; PRT; 496 AA.

AC Q93CP7

DT 01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)

DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

DE ALBF.

GN ALBF.

OS Xanthomonas albilineans.

OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;

OC Xanthomonas.

OX NCBI_TaxID=29447;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=XAL3;

RA Bostock J.M., Birch R.G.;

RT "A gene from Xanthomonas albilineans confers high-level albicidin resistance in Escherichia coli."

RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

RL EMBL; AF403709; AAL01877.1; --

DR InterPro; IPR003662; sub transporter.

DR Pfam; PF00083; sugar tr; 1.

SQ SEQUENCE 496 AA; 52709 MW; 4CBB0099FBE98867 CRC64;

Query Match 4.3%; Score 104.5; DB 2; Length 496;

Best Local Similarity 21.3%; Pred. No. 0.86;

Matches 92; Conservative 55; Mismatches 163; Indels 121; Gaps 21;

QY 41 LLSVGLLSVAACWFLPSI---TAA---AASWIITVLLCCSHKARCFILVFLSCGLREG 94

Db 41 LLDVGVINVA---LPSIOKNLGADQQLLEWIVAIYILFA-----LGLPLGLRGLGMDLG 91

QY 95 RNALTAAG-TGIVILG-----HVENIHFNFKGLDGMTC-----NLRA 131

Db 92 RKMEFTGVAGFILMSAFCAIAGNIHVLIARALQGLAAMLAPOVMAIAQTWFAKERA 151

QY 132 KFSFTHPLKKYIE-AIQWYGLATPLSVDFDLVSNW-----OTLAVSLFSP 178

Db 152 AAFSL-FGLVAGLASPAGPLVSGLLIHDAFG--VGMRAIFLINVPGLVTLIAAAIWP 208

QY 179 SHVLEAQL-NDKSGEVLSVLYQMATTEVLS--SLGQKLLAFAGLSVLLGTGLFMKRFL 235

Db 209 KVPAGAHGHNWVGIALAALLCLVFLPIEGRAYGWPLWCFAIAL-----255

QY 236 GFCGKYNENIYITROFQDERERHQRPCVLPNKERRKYVI---IPTFWTPKER 291

Db 256 -----GIPLVAFVQWRQAHARFALLPIYLMSHRDYILGALSUSVFYSA-----302

QY 292 NLGLFFPLIHLICWLVFAVDYLLYRLIFSQKQFQSLPGFEVHLKHGEKQGTQDII 351

Db 303 -LQGFLLVFIPLQOGLAYSALGTGATVTPFPVGVIAISM-----LARHVE-----347

QY 352 HDSSNFNISVFENPCPKPFLSETWVPLSVILLIIVMLGLSSILMQLKILVS-----405

Db 348 ---SLRAKIFSGACLMIASYL--ALWV-----IITRSGSLDPWTLTLLIIGLGGCI 396

QY 406 --ASFYPSVER 414

Db 397 TIASLQFQVMR 407

RESULT 13

Q9RNE1 PRELIMINARY; PRT; 320 AA.

ID Q9RNE1

AC Q9RNE1

DT 01-MAY-2000 (Tremblrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

DE Spore germination protein GerYB.

GN GERYS.

OS Bacillus anthracis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

OC Bacillaceae; Bacillus.

Result No.	Score	Query %		Length	DB	ID	Description
		Match	Length				
1	2448	100.0	470	22	AAE02638		Human dendritic ce
2	1799.5	73.5	352	22	AAB87357		Human gene 16 enco
3	1799.5	73.5	352	23	ABG55373		Human albumin fusi
4	1314	53.7	291	21	ABG55370		Human albumin fusi
5	1314	53.7	292	21	AAB34797		Human secreted pro
6	1309	53.5	257	22	AAB03917		Human gene 20 enco
7	1309	53.5	257	22	AAB87399		Human gene 16 enco
8	1309	53.5	257	22	AAB87390		Human gene 16 enco
9	1309	53.5	257	22	AAB87454		Human gene 16 enco
10	1309	53.5	257	23	ABG55371		Human albumin fusi

FT Domain , 295..313

ALIGNMENTS

FT /label= Transmembrane_domain_6
FT Modified-site 357...359
FT /note= "Asn is N-glycosylated"
FT Domain 379...398
FT /label= Transmembrane_domain_7
XX
PN WO200136463-A2.
XX
PD 25-MAY-2001.
XX
XX 15-NOV-2000; 2000WO-US31167.
XX PF
XX 15-NOV-1999; 99US-0439735.
XX PR
XX (SCHE) SCHERING CORP.
PA
XX Zlot CH, Adema GJ, Figdor C, Phillips JH;
XX
XX WPI: 2001-367562/38.
DR N-PSDB; AAD06864.
XX
XX New mammalian proteins designated dendritic cell specific transmembrane
PT protein and DNAX surface protein and the nucleic acids encoding the
PT polypeptides -
XX
XX Claim 2; Page 14; 96pp; English.
XX
XX The present sequence is human dendritic cell specific transmembrane
CC protein (DC-STAMP). The DC-STAMP is useful in forensic sciences,
CC e.g. to distinguish rodent from human or as a marker to distinguish
CC between different cells exhibiting differential expression or
CC modification patterns. The DC-STAMP and its antibodies are used for
CC preparing kits for use in molecular biology, immunology or physiology,
CC and in treatment of conditions associated with abnormal physiology or
CC development, e.g. allergic conditions (asthma), cancer and autoimmune
CC diseases (diabetes mellitus). Drug screening using DC-STAMP or its
CC fragments can be performed to identify compounds having binding affinity
CC to or other relevant isolation effects on the function of DC-STAMP.
CC The DC-STAMP can also be used in diagnostic kits and methods for
CC detecting the presence of another DC-STAMP or binding partner.
XX
SQ Sequence 470 AA;

Query Match 100.0%; Score 2448; DB 22; Length 470;
Best Local Similarity 100.0%; Pred. No. 4.2e-247;
Matches 470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGIIWTSQTDIFLSLWEIYVSPSPGWMDFIOHLGVCCVLVALISVGLLSVAACWFLPSIIA 60
DB 1 MGIIWTSQTDIFLSLWEIYVSPSPGWMDFIOHLGVCCVLVALISVGLLSVAACWFLPSIIA 60

QY 61 AAASMIITCVLLCCSKHARCFILLVFLSCGLREGRNALIAAGTGVILGHVENIHFNFKG 120
DB 61 AAASMIITCVLLCCSKHARCFILLVFLSCGLREGRNALIAAGTGVILGHVENIHFNFKG 120

QY 121 LLDGMCNLRKASFSTHFPLLKKYIEAIOIYIGLATPLSVFDDLVSNWOTLAVSFSPSH 180
DB 121 LLDGMCNLRKASFSTHFPLLKKYIEAIOIYIGLATPLSVFDDLVSNWOTLAVSFSPSH 180

QY 181 VLEAQLNDSKGEVLVLYQMATTEVLSLQKLLAFAGLSVLVLTGLFMKRFGLPCGW 240
DB 181 VLEAQLNDSKGEVLVLYQMATTEVLSLQKLLAFAGLSVLVLTGLFMKRFGLPCGW 240

QY 241 KYENIYITQFQFDEREHOORPCVLPLNKEERRYVIPTFWPTPKRKNLGLFFLPI 300
DB 241 KYENIYITQFQFDEREHOORPCVLPLNKEERRYVIPTFWPTPKRKNLGLFFLPI 300

QY 301 LIHLCIWLFAADVLLYLFISVSKQFSLPGFEVHLKHGEKQGTQDIHDSFNISV 360
DB 301 LIHLCIWLFAADVLLYLFISVSKQFSLPGFEVHLKHGEKQGTQDIHDSFNISV 360

QY 361 FEPNCIPKPKFLLSETWVPLSVLLILVMLGLSSILMQLKILVSASFYPSVERKRIQYL 420

Db 361 FEPNCIPKPKFLLSETWVPLSVLLILVMLGLSSILMQLKILVSASFYPSVERKRIQYL 420
QY 421 HAKLLKSGSKQPLGEVGRRLSLYLTKEHFWLPVLKMKRKQMDMASADKS 470
Db 421 HAKLLKSGSKQPLGEVGRRLSLYLTKEHFWLPVLKMKRKQMDMASADKS 470

RESULT 2
AAB87357
ID AAB87357 standard; Protein; 352 AA.
XX
AC AAB87357;
XX
DT 22-MAY-2001 (first entry)
XX
DE Human gene 16 encoded secreted protein HMADJ14, SEQ ID NO:98.
XX
XX Human; secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angiogenic disorder; kidney disorder;
KW gastrointestinal disorder; infection; pregnancy-related disorder;
KW endocrine disorder; infection; wound healing; vulneryary;
KW cell culture; chemotaxis; food additive;
KW binding partner identification.
XX
XX Homo sapiens.
OS
PN WO200118022-A1.
PD
PD 15-MAR-2001.
PF 31-AUG-2000; 2000WO-US24008.
XX
PR 03-SEP-1999; 99US-0152315.
PR 03-SEP-1999; 99US-0152317.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Ni J, Baker KP, Birse CE, Fiscella M, Komatsoulis CA, Rosen CA;
PI Soppet DR, Young PE, Ebner R, Duan DR, Olsen HS, Lafleur DW;
PI Moore PA, Shi Y, Wei Y, Florence KA;
XX
DR WPI: 2001-203081/20.
DR N-PSDB; AAF91873.
XX
PT Nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT Parkinson's diseases and cancers -
XX
XX Claim 11; Page 544-545; 607pp; English.
XX
CC AAF91858-AAF91929 represent cDNAs corresponding to 52 human secreted
CC protein genes, and AAB87342-AAB87413 represent the proteins they encode.
CC AAB87414-AAB87454 represent human secreted protein fragments. The genes
CC and their corresponding secreted proteins are useful for preventing,
CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Pathological conditions can be diagnosed by determining the
CC amount of the new protein in a sample or by determining the presence of
CC mutations in the new genes. Specific uses are described for each of the
CC 52 genes, based on the tissues in which they are most highly expressed,
CC and include developing products for the diagnosis or treatment of
CC proliferative disorders, cancer, tumours, foetal and developmental
CC abnormalities, haematopoietic disorders, diseases of the immune system,
CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
CC allergies, neurological disorders (e.g., Alzheimer's disease,
CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC cardiovascular disorders, angiogenic disorders, kidney disorders,
CC gastrointestinal disorders, pregnancy-related disorders, endocrine

disorders, and infections. The proteins can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues, to identify their cognate ligands or binding partners, and in chemotaxis, and can be used as a food additive or preservative to modify storage properties. Antibodies specific for a protein of the invention can be used in alleviating symptoms associated with the disorders mentioned above, and in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The present sequence represents a human secreted protein of the invention.

XX SQ Sequence 352 AA;

Query Match 73.5%; Score 1799.5; DB 22; Length 352;

Best Local Similarity 96.4%; Pred. No. 2e-179;

Matches 344; Conservative 2; Mismatches 4; Indels 7; Gaps 1;

QY 27 MDFIQHLGVCCCLVALISVGLLSVAACWFLPSIIAAAASWIIITCVLLCCSKHARCFILLVF 86

DB 1 MDFIQHLGVCCCLVALISVGLLSVAACWFLPSIIAAAASWIIITCVLLCCSKHARCFILLVF 60

QY 87 LSCGLREGNALLAAGTGIVILGHVENIFHNFKGLLDGWTNLRKAKSFIHFFLLKKYIE 146

DB 61 LSCGLREGNALLAAGTGIVILGHVENIFHNFKGLLDGWTNLRKAKSFIHFFLLKKYIE 120

QY 147 AIQWYIGLATPLSVFDDLVSNQNTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 206

DB 121 AIQWYIGLATPLSVFDDLVSNQNTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 180

QY 207 LSSIQOKLLAFAGLSVLVLTGTLFMKRFGLPCGWKYNIIYITRQVQFDERERHQORPCV 266

DB 181 LSSIQOKLLAFAGLSVLVLTGTLFMKRFGLPCGWKYNIIYITRQVQFDERERHQORPCV 240

QY 267 LPLNKEERRKVIITFTWPTPKERNKGLFPLILHLICWLVFAADVLYLRLFSVSK 326

DB 241 LPLNKEERRKVIITFTWPTPKERNKGLFPLILHLICWLVFAADVLYLRLFSVSK 300

QY 327 QFQSLPGFEVHLKLHGEKQGTQDIIHDSFNISVFEPNCIPKPKFLLSETWVPLSVI 383

DB 301 QFQSLPGFEVHLKLHGEKQGTQDIIHDSFNISVFEPNCIPKPKFLLSETWVPLSVI 350

RESULT 3

ABG65373

ID ABG65373 standard; Protein; 352 AA.

XX AC ABG65373;

XX DT 27-AUG-2002 (first entry)

XX DE Human albumin fusion protein #2048.

XX KW Albumin fusion protein; therapeutic protein X; human albumin; HA;

XX KW human serum albumin; HSA; cancer; reproductive disorder;

XX KW digestive disorder; immune disorder; endocrine disorder;

XX KW haematopoietic disorder; neural disorder; connective disorder;

XX KW cytosolic; antifertility; antinflammatory; antiulcer;

XX KW immunomodulator; anti-HIV; antidiabetic; haemostatic; neutropic;

XX KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;

XX KW osteopathic; antiarthritic.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200177137-A1.

XX PD 18-OCT-2001.

XX PF 12-APR-2001; 2001WO-US11989.

XX PR 12-APR-2000; 2000US-229358P.

XX PR 25-APR-2000; 2000US-199384P.

PR 21-DEC-2000; 2000US-256931P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Haseltine WA;

PI WPI; 2002-010886/01.

DR New fusion protein for treating disease e.g. diabetes comprises an

XX PT albumin fused to a therapeutic protein -

XX PS Claim 1; Page 1954-1955; 2102pp; English.

XX CC The present invention relates to albumin fusion proteins comprising a

CC therapeutic protein X and human albumin (HA, also known as human serum

CC albumin, HSA). The proteins are useful for treating a disease or

CC disorder that may be modulated by therapeutic protein X. The albumin

CC extends the shelf-life of protein X, and may increase its biological

CC in vitro/in vivo activity. The protein is useful for treating and

CC diagnosing disorders such as cancer, reproductive disorders, digestive

CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders

CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders

CC (e.g. diabetes), haematopoietic disorders, neural disorders

CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,

CC encephalomyelitis, meningitis, schizophrenia), and connective disorders

CC (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin

CC fusion proteins of the invention.

XX SQ Sequence 352 AA;

Query Match 73.5%; Score 1799.5; DB 23; Length 352;

Best Local Similarity 96.4%; Pred. No. 2e-179;

Matches 344; Conservative 2; Mismatches 4; Indels 7; Gaps 1;

QY 27 MDFIQHLGVCCCLVALISVGLLSVAACWFLPSIIAAAASWIIITCVLLCCSKHARCFILLVF 86

DB 1 MDFIQHLGVCCCLVALISVGLLSVAACWFLPSIIAAAASWIIITCVLLCCSKHARCFILLVF 60

QY 87 LSCGLREGNALLAAGTGIVILGHVENIFHNFKGLLDGWTNLRKAKSFIHFFLLKKYIE 146

DB 61 LSCGLREGNALLAAGTGIVILGHVENIFHNFKGLLDGWTNLRKAKSFIHFFLLKKYIE 120

QY 147 AIQWYIGLATPLSVFDDLVSNQNTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 206

DB 121 AIQWYIGLATPLSVFDDLVSNQNTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 180

QY 207 LSSIQOKLLAFAGLSVLVLTGTLFMKRFGLPCGWKYNIIYITRQVQFDERERHQORPCV 266

DB 181 LSSIQOKLLAFAGLSVLVLTGTLFMKRFGLPCGWKYNIIYITRQVQFDERERHQORPCV 240

QY 267 LPLNKEERRKVIITFTWPTPKERNKGLFPLILHLICWLVFAADVLYLRLFSVSK 326

DB 241 LPLNKEERRKVIITFTWPTPKERNKGLFPLILHLICWLVFAADVLYLRLFSVSK 300

QY 327 QFQSLPGFEVHLKLHGEKQGTQDIIHDSFNISVFEPNCIPKPKFLLSETWVPLSVI 383

DB 301 QFQSLPGFEVHLKLHGEKQGTQDIIHDSFNISVFEPNCIPKPKFLLSETWVPLSVI 350

RESULT 4

ABG65370

ID ABG65370 standard; Protein; 291 AA.

XX AC ABG65370;

XX DT 27-AUG-2002 (first entry)

XX DE Human albumin fusion protein #2045.

XX KW Albumin fusion protein; therapeutic protein X; human albumin; HA;

XX KW human serum albumin; HSA; cancer; reproductive disorder;

XX KW digestive disorder; immune disorder; endocrine disorder;

XX KW haematopoietic disorder; neural disorder; connective disorder;

KW cytostatic; antiinfertility; antiinflammatory; antiulcer;
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; neutropic;
KW neutroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW osteopathic; antiarthritic.
OS Homo sapiens.
OS Synthetic.
XX WO200177137-A1.
XX 18-OCT-2001.
XX 12-APR-2001; 2001WO-US11988.
XX 12-APR-2000; 2000US-229358P.
XX 25-APR-2000; 2000US-199384P.
XX 21-DEC-2000; 2000US-256931P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Haseltine WA;
PI WPI; 2002-010886/01.
XX
XX New fusion protein for treating disease e.g. diabetes comprises an
PT albumin fused to a therapeutic protein -
PT
XX
XX Claim 1; Page 1951-1952; 2102pp; English.
XX
XX The present invention relates to albumin fusion proteins comprising a
CC therapeutic protein X and human albumin (HA, also known as human serum,
CC albumin, HSA). The proteins are useful for treating a disease or
CC disorder that may be modulated by therapeutic protein X. The albumin
CC extends the shelf-life of protein X, and may increase its biological
CC in vitro/in vivo activity. The protein is useful for treating and
CC diagnosing disorders such as cancer, reproductive disorders, digestive
CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
CC (e.g. diabetes), haematopoietic disorders, neural disorders
CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
CC encephalomyelitis, meningitis, schizophrenia), and connective disorders
CC (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin
CC fusion proteins of the invention.
XX
SQ Sequence 291 AA;
Query Match 53.7%; Score 1314; DB 23; Length 291;
Best Local Similarity 92.5%; Pred. No. 9.9e-129;
Matches 259; Conservative 4; Mismatches 5; Indels 12; Gaps 3;
QY 27 MDPIQHLGVCCCLVALISVGLLSVAACWFLPSIIAAASWIITCVLLCCSKHARCFTLLVP 86
DB 1 MDPIQHLGVCCCLVALISVGLLSVAACWFLPSIIAAASWIITCVLLCCSKHARCFTLLVP 60
QY 87 LSCGLREGNALLIAGTGIVILGHVENIIFHNFKGLDGMTCNLRAKFSIHFPLLKKYIE 146
DB 61 LSCGLREGNALLIAGTGIVILGHVENIIFHNFKGLDGMTCNLRAKFSIHFPLLKKYIE 120
QY 147 ATQWLYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGVLSVLYQMATTVEV 206
DB 121 ATQWLYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGVLSVLYQMATTVEV 180
QY 207 LSSLGQKLLAFAGLSVLVLTGTLFMKRFGLGPCWKYENIYITROFVQFDERERHQPCV 266
DB 181 LSSLGQKLLAFAGLSVLVLTGTLFMKRFGLGPCWKYENIYITROFVQFDERERHQPCV 240
QY 267 LPLNKEERKYYIITFWPTPKERKNLGLFFLPILHLCI 306
DB 241 LPLNKEERK-----NKLKILSM-IIP-LIYLCL 268
RESULT 5
AAB34797

ID AAB34797 standard; Protein; 292 AA.
AC AAB34797;
XX 26-JAN-2001 (first entry)
XX Human secreted protein sequence encoded by gene 25 SEQ ID NO:85.
XX
XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
KW neutropic; neutroprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiant; gene therapy;
KW cancer; immune disorder; cardiovascular disorder; wound healing;
KW neurological disease; infectious disease; chromosome identification.
XX
XX Homo sapiens.
XX WO2000058356-A1.
XX 05-OCT-2000.
XX 22-MAR-2000; 2000WO-US07535.
XX 26-MAR-1999; 99US-0126511.
XX 17-DEC-1999; 99US-0172413.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM, Komatsoulis G;
PI WPI; 2000-594639/56.
XX N-PSDB; AAC59990.
XX Fifty nucleic acid molecules encoding human secreted proteins, useful
PT in the prevention, treatment and diagnosis of cancer, immune disorders,
PT cardiovascular disorders and neurological diseases -
XX
XX Claim 1; Page 385-386; 425pp; English.
XX The polynucleotide sequences given in AAC59966 to AAC60015 encode the
CC human secreted proteins given in AAB34773 to AAB34822. AAB34823 to
CC AAB34852 represent human secreted polypeptide sequences and proteins
CC homologous to them, which are given in the exemplification of the present
CC invention. Human secreted proteins have activities based on the tissues
CC and cells the genes are expressed in. Examples of activities include:
CC cytostatic; immunosuppressive; neutropic; neutroprotective; antiviral;
CC antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
CC vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic; and
CC cardiant. The polynucleotides and polypeptides are useful for
CC preventing, treating or ameliorating a medical condition in e.g. humans,
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. The polynucleotides are
CC useful for chromosome identification. They are also useful as probes for
CC diagnosing a disorder related to the female reproductive system,
CC particularly breast and/or ovary cancer. They are also useful in the gene
CC therapy of breast and ovarian cancer. Nucleic acids, protein, antibodies,
CC agonists and antagonists from the present invention are useful in the
CC diagnosis, treatment and prevention of cancer, immune disorders,
CC cardiovascular disorders, wound healing, neurological diseases and
CC infectious disease. AAC59957 to AAC59965 and AAB34772 represents sequence
CC used in the exemplification of the present invention.
XX
SQ Sequence 292 AA;
Query Match 53.7%; Score 1314; DB 21; Length 292;
Best Local Similarity 92.5%; Pred. No. 1e-128;
Matches 259; Conservative 4; Mismatches 5; Indels 12; Gaps 3;
QY 27 MDPIQHLGVCCCLVALISVGLLSVAACWFLPSIIAAASWIITCVLLCCSKHARCFTLLVP 86
DB 1 MDPIQHLGVCCCLVALISVGLLSVAACWFLPSIIAAASWIITCVLLCCSKHARCFTLLVP 60

QY 87 LSCGLREGNALLAAGTGIVILGHVENIFHNFKGLLDGWTNLRKASFSIHFPPLKKYIE 146
 Db 61 LSCGLREGNALLAAGTGIVILGHVENIFHNFKGLLDGWTNLRKASFSIHFPPLKKYIE 120
 QY 147 AIQWYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 206
 Db 121 AIQWYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 180
 QY 207 LSSLGQKLLAFAGLSVLVLTGTLFMKRFGLGPCWKYENIYITRQVQFDERERHQORPCV 266
 Db 181 LSSLGQKLLAFAGLSVLVLTGTLFMKRFGLGPCWKYENIYITRQVQFDERERHQORPCV 240
 QY 267 LPLNKEERRKYYIIPFWPTPKERNGLGFLPILIHLCI 306
 Db 241 LPLNKEERRK-----NRELKILSM-ILP-LIYLCL 268

RESULT 6
 AAE03917
 ID AAE03917 standard; Protein; 257 AA.
 XX AC
 XX AA03917;
 XX DT
 XX 09-AUG-2001 (first entry)
 XX DE Human gene 20 encoded secreted protein HMADJ14, SEQ ID NO:80.
 XX KW Human; secreted protein; proliferative disorder; cancer; tumour;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiogenic disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder;
 KW endocrine disorder; infection; wound healing; vulnery;
 KW cell culture; chemotaxis; food additive; gene therapy;
 KW binding partner identification.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT Peptide 1..26
 FT /label= signal_peptide
 FT Protein 27..257
 FT /note= "Mature secreted protein"
 XX W0200077022-A1.
 XX PN
 XX PD 21-DEC-2000.
 XX PF 01-JUN-2000; 2000WO-US15136.
 XX PR 11-JUN-1999; 99US-0138629.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Rosen CA, Ruben SM, Komatsoulis GA;
 XX WPI; 2001-367020/38.
 XX DR N-PSDB; AAD08364.
 XX PT Nucleic acids encoding 50 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating diseases, e.g. Parkinson's
 PT disease, botulism, cancers and Scimitar syndrome -
 XX Claim 11; Page 535-536; 614pp; English.
 XX PS
 CC AAD08345-AAD08394 represent cDNAs corresponding to 50 human secreted
 CC protein genes and AAE03988-AAE03947 represent the proteins they encode.
 CC AAE03948-AAE03996 represent human secreted protein fragments or variants.
 CC The genes and their secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene

CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 50 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC proliferative disorders, cancer, tumours, foetal and developmental
 CC abnormalities, haematopoietic disorders, diseases of the immune system,
 CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
 CC allergies, neurological disorders (e.g., Alzheimer's disease,
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiogenic disorders, kidney disorders,
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine
 CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties.
 CC Antibodies specific for a protein of the invention can be used in
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein of the invention.
 XX

Sequence 257 AA;

Query Match 53.5%; Score 1309; DB 22; Length 257;
 Best Local Similarity 99.2%; Pred. No. 2.8e-128;
 Matches 250; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 27 MDFIOHLGVCCILVALISVGLLSVAACWFLPSIIAAAASWIIITCVLLCCSKHARCFILLVF 86
 Db 1 MDFIOHLGVCCILVALISVGLLSVAACWFLPSIIAAAASWIIITCVLLCCSKHARCFILLVF 60
 QY 87 LSCGLREGNALLAAGTGIVILGHVENIFHNFKGLLDGWTNLRKASFSIHFPPLKKYIE 146
 Db 61 LSCGLREGNALLAAGTGIVILGHVENIFHNFKGLLDGWTNLRKASFSIHFPPLKKYIE 120
 QY 147 AIQWYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 206
 Db 121 AIQWYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 180
 QY 207 LSSLGQKLLAFAGLSVLVLTGTLFMKRFGLGPCWKYENIYITRQVQFDERERHQORPCV 266
 Db 181 LSSLGQKLLAFAGLSVLVLTGTLFMKRFGLGPCWKYENIYITRQVQFDERERHQORPCV 240
 QY 267 LPLNKEERRKYYV 278
 Db 241 LPLNKEERRKFI 252

RESULT 7

AAE03917
 ID AAE03917 standard; Protein; 257 AA.
 XX AC
 XX AA03917;
 XX DT
 XX 22-MAY-2001 (first entry)
 XX DE Human gene 16 encoded secreted protein HMADJ14, SEQ ID NO:140.

Human; secreted protein; proliferative disorder; cancer; tumour;
 foetal abnormality; developmental abnormality; haematopoietic disorder;
 immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiogenic disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder;
 KW endocrine disorder; infection; wound healing; vulnery;
 KW cell culture; chemotaxis; food additive;
 KW binding partner identification.


```
XX OS Homo sapiens.
XX PN WO200118022-A1.
XX PD 15-MAR-2001.
XX PF 31-AUG-2000; 2000WO-US24008.
XX PR 03-SEP-1999; 99US-0152315.
XX PR 03-SEP-1999; 99US-0152317.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Ni J, Baker KP, Birse CE, Fiscella M, Komatsoulis GA, Rosen CA;
PI Soppet DR, Young PE, Ebner R, Duan DR, Olsen HS, Lafleur DW;
PI Moore PA, Shi Y, Wei Y, Florence KA;
XX WPI; 2001-203081/20.
XX DR N-PSDB; AAF91915.
XX PT Nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT Parkinson's diseases and cancers -
XX PS Claim 11; Page 567-568; 607pp; English.
XX CC AAF91858-AAF91929 represent cDNAs corresponding to 52 human secreted
CC protein genes, and AAB87342-AAB87413 represent the proteins they encode.
CC AAB87414-AAB87454 represent human secreted protein fragments. The genes
CC and their corresponding secreted proteins are useful for preventing,
CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Pathological conditions can be diagnosed by determining the
CC amount of the new protein in a sample or by determining the presence of
CC mutations in the new genes. Specific uses are described for each of the
CC 52 genes, based on the tissues in which they are most highly expressed,
CC and include developing products for the diagnosis or treatment of
CC proliferative disorders, cancer, tumours, foetal and developmental
CC abnormalities, haematopoietic disorders, diseases of the immune system,
CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
CC allergies, neurological disorders (e.g., Alzheimer's disease,
CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC cardiovascular disorders, angiogenic disorders, kidney disorders,
CC gastrointestinal disorders, pregnancy-related disorders, endocrine
CC disorders, and infections. The proteins can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin aging due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues, to identify their
CC cognate ligands or binding partners, and in chemotaxis, and can be used
CC as a food additive or preservative to modify storage properties.
CC Antibodies specific for a protein of the invention can be used in
CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunosorbent assay (ELISA). The present sequence represents a human
CC secreted protein of the invention.
XX SQ Sequence 257 AA;
Query Match 53.5%; Score 1309; DB 22; Length 257;
Best Local Similarity 99.2%; Pred. No. 2.8e-128;
Matches 250; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 27 MDFOHGLGVCCLVALISVGLLSVACWFLPSIAAASWITCVLLCCSKHARCFILLVF 86
DB 1 MDFOHGLGVCCLVALISVGLLSVACWFLPSIAAASWITCVLLCCSKHARCFILLVF 60
QY 87 LSCGLREGNALLIAGTGIVILGHVENI FHNFKGLDGMTCNLRKSFSIHPFLKKYTE 146
DB 61 LSCGLREGNALLIAGTGIVILGHVENI FHNFKGLDGMTCNLRKSFSIHPFLKKYTE 120
QY 147 AIQWIYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 206
,
```

```
DB 121 AIQWIYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 180
QY 207 LSSLGQKLIAPAGLSLVLLGTGLPMKRFLGPCGWKYENIYITRQFVQFDERRHRQPCV 266
DB 181 LSSLGQKLIAPAGLSLVLLGTGLPMKRFLGPCGWKYENIYITRQFVQFDERRHRQPCV 240
QY 267 LPLNKEERRKYV 278
DB 241 LPLNKEERRKFI 252
RESULT 8
AAB87400
ID AAB87400 standard; Protein; 257 AA.
XX AC AAB87400;
XX DT 22-MAY-2001 (first entry)
XX DE Human gene 16 encoded secreted protein HMAJ14, SEQ ID NO:141.
XX KW Human; secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angiogenic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
KW endocrine disorder; infection; wound healing; vulnary;
KW cell culture; chemotaxis; food additive;
XX binding partner identification.
XX OS Homo sapiens.
XX PN WO200118022-A1.
XX PD 15-MAR-2001.
XX PF 31-AUG-2000; 2000WO-US24008.
XX PR 03-SEP-1999; 99US-0152315.
XX PR 03-SEP-1999; 99US-0152317.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Ni J, Baker KP, Birse CE, Fiscella M, Komatsoulis GA, Rosen CA;
PI Soppet DR, Young PE, Ebner R, Duan DR, Olsen HS, Lafleur DW;
PI Moore PA, Shi Y, Wei Y, Florence KA;
XX WPI; 2001-203081/20.
XX DR N-PSDB; AAF91916.
XX PT Nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT Parkinson's diseases and cancers -
XX PS Claim 11; Page 568-569; 607pp; English.
XX CC AAF91858-AAF91929 represent cDNAs corresponding to 52 human secreted
CC protein genes, and AAB87342-AAB87413 represent the proteins they encode.
CC AAB87414-AAB87454 represent human secreted protein fragments. The genes
CC and their corresponding secreted proteins are useful for preventing,
CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Pathological conditions can be diagnosed by determining the
CC amount of the new protein in a sample or by determining the presence of
CC mutations in the new genes. Specific uses are described for each of the
CC 52 genes, based on the tissues in which they are most highly expressed,
CC and include developing products for the diagnosis or treatment of
CC proliferative disorders, cancer, tumours, foetal and developmental
CC abnormalities, haematopoietic disorders, diseases of the immune system,
CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
CC allergies, neurological disorders (e.g., Alzheimer's disease,
```

CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiogenic disorders, kidney disorders,
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine
 CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties.
 CC Antibodies specific for a protein of the invention can be used in
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein of the invention.
 XX
 SQ Sequence 257 AA;
 Query Match 53.5%; Score 1309; DB 22; Length 257;
 Best Local Similarity 99.2%; Pred. No. 2.8e-128;
 Matches 250; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 27 MDFIQLHGVCCILVALISVGLLSVAACWFLPSIIIAAASWIIITCVLLCCSKHARCFILLVF 86
 DB 1 MDFIQLHGVCCILVALISVGLLSVAACWFLPSIIIAAASWIIITCVLLCCSKHARCFILLVF 60
 QY 87 LSCGLREGNALLIAAGTGIVILGHVENIFHNFKGLDGMTCNLRKSFHFFLLKKYIE 146
 DB 61 LSCGLREGNALLIAAGTGIVILGHVENIFHNFKGLDGMTCNLRKSFHFFLLKKYIE 120
 QY 147 AIQWYIGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 206
 DB 121 AIQWYIGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 180
 QY 207 LSSLGQKLIAPAGLSVLILGTGLFMKRFGLGPCGKYNIIYITRFQVQFDERERHQRPCV 266
 DB 181 LSSLGQKLIAPAGLSVLILGTGLFMKRFGLGPCGKYNIIYITRFQVQFDERERHQRPCV 240
 QY 267 LPLNKEERRKYV 278
 DB 241 LPLNKEERRKFI 252
 RESULT 9
 AAB87454
 ID AAB87454 standard; Protein; 257 AA.
 XX
 AC AAB87454;
 XX
 DT 22-MAY-2001 (first entry)
 XX
 DE Human gene 16 encoded secreted protein fragment.
 XX
 KW Human; secreted protein; proliferative disorder; cancer; tumour;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiogenic disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder;
 KW endocrine disorder; infection; wound healing; vulnery;
 KW cell culture; chemotaxis; food additive;
 KW binding partner identification.
 XX
 OS Homo sapiens.
 XX
 PN WO200118022-A1.
 XX
 XX 15-MAR-2001.
 PD
 XX 31-AUG-2000; 2000WO-US24008.
 PF
 XX

PR 03-SEP-1999; 99US-0152315.
 PR 03-SEP-1999; 99US-0152317.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ni J. Baker KP, Birse CE, Fiscella M, Komatsoulis GA, Rosen CA;
 PI Soppet DR, Young PE, Ebner R, Duan DR, Olsen HS, Lafleur DW;
 PI Moore PA, Shi Y, Wei Y, Florence KA;
 XX
 DR WPI; 2001-203081/20.
 XX
 XX Nucleic acid molecules encoding human secreted proteins, used in
 PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
 PT Parkinson's diseases and cancers -
 XX
 PS Disclosure; Page 55; 607pp; English.
 XX
 CC AAF91858-AAF91929 represent cDNAs corresponding to 52 human secreted
 CC protein genes, and AAB87342-AAB87413 represent the proteins they encode.
 CC AAB87414-AAB87454 represent human secreted protein fragments. The genes
 CC and their corresponding secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 52 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC proliferative disorders, cancer, tumours, foetal and developmental
 CC abnormalities, haematopoietic disorders, diseases of the immune system,
 CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
 CC allergies, neurological disorders (e.g., Alzheimer's disease,
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiogenic disorders, kidney disorders,
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine
 CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties.
 CC Antibodies specific for a protein of the invention can be used in
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein fragment referred to in the disclosure of the invention.
 XX
 SQ Sequence 257 AA;
 Query Match 53.5%; Score 1309; DB 22; Length 257;
 Best Local Similarity 99.2%; Pred. No. 2.8e-128;
 Matches 250; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 27 MDFIQLHGVCCILVALISVGLLSVAACWFLPSIIIAAASWIIITCVLLCCSKHARCFILLVF 86
 DB 1 MDFIQLHGVCCILVALISVGLLSVAACWFLPSIIIAAASWIIITCVLLCCSKHARCFILLVF 60
 QY 87 LSCGLREGNALLIAAGTGIVILGHVENIFHNFKGLDGMTCNLRKSFHFFLLKKYIE 146
 DB 61 LSCGLREGNALLIAAGTGIVILGHVENIFHNFKGLDGMTCNLRKSFHFFLLKKYIE 120
 QY 147 AIQWYIGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 206
 DB 121 AIQWYIGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 180
 QY 207 LSSLGQKLIAPAGLSVLILGTGLFMKRFGLGPCGKYNIIYITRFQVQFDERERHQRPCV 266
 DB 181 LSSLGQKLIAPAGLSVLILGTGLFMKRFGLGPCGKYNIIYITRFQVQFDERERHQRPCV 240
 QY 267 LPLNKEERRKYV 278
 DB 241 LPLNKEERRKFI 252

```

RESULT 10
ABG65371
ID ABG65371 standard; Protein; 257 AA.
XX
AC ABG65371;
XX
DT 27-AUG-2002 (first entry)
XX
DE Human albumin fusion protein #2046.
XX
KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW human serum albumin; HSA; cancer; reproductive disorder;
KW digestive disorder; immune disorder; endocrine disorder;
KW haematopoietic disorder; neural disorder; connective disorder;
KW cytostatic; antiinfertility; antiinflammatory; antiulcer;
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW osteopathic; antiarthritic.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200177137-A1.
XX
PD 18-OCT-2001.
XX
PF 12-APR-2001; 2001WO-US11988.
XX
PR 12-APR-2000; 2000US-229358P.
PR 25-APR-2000; 2000US-199384P.
PR 21-DEC-2000; 2000US-256931P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Haseltine WA;
XX
WPI; 2002-010886/01.
XX
New fusion protein for treating disease e.g. diabetes comprises an
albumin fused to a therapeutic protein -
Claim 1; Page 1952-1953; 2102pp; English.
XX
The present invention relates to albumin fusion proteins comprising a
therapeutic protein X and human albumin (HA, also known as human serum
albumin, HSA). The proteins are useful for treating a disease or
disorder that may be modulated by therapeutic protein X. The albumin
extends the shelf-life of protein X, and may increase its biological
in vitro/in vivo activity. The protein is useful for treating and
diagnosing disorders such as cancer, reproductive disorders, digestive
disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
(e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
(e.g. diabetes), haematopoietic disorders, neural disorders
(e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
encephalomyelitis, meningitis, schizophrenia), and connective disorders
(e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin
fusion proteins of the invention.
XX
SQ Sequence 257 AA;

Query Match
Best Local Similarity 99.2%; Score 1309; DB 23; Length 257;
Matches 250; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

27 MDFIQHLGVCCVLVALISVGLLSVAACWFLPSIIIAAASWITCVLLCCSKHARCFILLVF 86
1 MDFIQHLGVCCVLVALISVGLLSVAACWFLPSIIIAAASWITCVLLCCSKHARCFILLVF 60
87 LSCGLREGCNALIAAGTGIVLGHVENIFHNFKGLLDGWTCLNRAKSFHIFPLKKYIE 146
61 LSCGLREGCNALIAAGTGIVLGHVENIFHNFKGLLDGWTCLNRAKSFHIFPLKKYIE 120

```

```
Query Match      53.5%; Score 1309; DB 23; Length 257;
Best Local Similarity 99.2%; Pred. No. 2.8e-128;
Matches 250; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 27 MDFIQHLGVCCLVALLISVGLLSVAACWFLPSIIAAAASWIIITCVLLCCSKHARCFILLVVF 86
Db 1 MDFIQHLGVCCLVALLISVGLLSVAACWFLPSIIAAAASWIIITCVLLCCSKHARCFILLVVF 60

QY 87 LSCGLREGNALLAAGTGIVILGHVENIFHNPKGLLDGTCNLRKAKSFIHPPLKKYIE 146
Db 61 LSCGLREGNALLAAGTGIVILGHVENIFHNPKGLLDGTCNLRKAKSFIHPPLKKYIE 120

QY 147 AIQWIYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 206
Db 121 AIQWIYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 180

QY 207 LSSLGQKLLAFAGLSVLVLTGTLFMRKFLGPGCWKYENIYITRQVFQDERERHQORPCV 266
Db 181 LSSLGQKLLAFAGLSVLVLTGTLFMRKFLGPGCWKYENIYITRQVFQDERERHQORPCV 240

QY 267 LPLNKEERKYYV 278
Db 241 LPLNKEERKFI 252

RESULT 12
AAB87401
ID AAB87401 standard; Protein; 291 AA.
XX
AC AAB87401;
XX
DT 22-MAY-2001 (first entry)
DE Human gene 16 encoded secreted protein HMADJ74, SEQ ID NO:142.
DE
XX Human; secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angiogenic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
KW endocrine disorder; infection; wound healing; vulnery;
KW cell culture; chemotaxis; food additive;
KW binding partner identification.
XX Homo sapiens.
OS
XX WO200118022-A1.
XX
XX 15-MAR-2001.
XX
XX 31-AUG-2000; 2000WO-US24008.
XX
XX 03-SEP-1999; 99US-0152315.
XX
XX 03-SEP-1999; 99US-0152317.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ni J, Baker KP, Birse CE, Fiscella M, Komatsoulis GA, Rosen CA;
XX Soppet DR, Young PE, Ebner R, Duan DR, Olsen HS, Lafleur DW;
XX Moore PA, Shi Y, Wei Y, Florence KA;
XX
XX WPI; 2001-203081/20.
XX
XX N-PSDB; AAF91917.
XX
XX Nucleic acid molecules encoding human secreted proteins, used in
XX preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
XX Parkinson's diseases and cancers -
XX
XX Claim 11; Page 569-570; 607pp; English.
XX
```

```
CC AAF91858-AAF91929 represent cDNAs corresponding to 52 human secreted
CC protein genes, and AAB87342-AAB87413 represent the proteins they encode.
CC AAB87414-AAB87454 represent human secreted protein fragments. The genes
CC and their corresponding secreted proteins are useful for preventing,
CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Pathological conditions can be diagnosed by determining the
CC amount of the new protein in a sample or by determining the presence of
CC mutations in the new genes. Specific uses are described for each of the
CC 52 genes, based on the tissues in which they are most highly expressed,
CC and include developing products for the diagnosis or treatment of
CC proliferative disorders, cancer, tumours, foetal and developmental
CC abnormalities, haematopoietic disorders, diseases of the immune system,
CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
CC allergies, neurological disorders (e.g., Alzheimer's disease,
CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC cardiovascular disorders, angiogenic disorders, kidney disorders,
CC gastrointestinal disorders, pregnancy-related disorders, endocrine
CC disorders, and infections. The proteins can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin aging due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues, to identify their
CC cognate ligands or binding partners, and in chemotaxis, and can be used
CC as a food additive or preservative to modify storage properties.
CC Antibodies specific for a protein of the invention can be used in
CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunoassay (ELISA). The present sequence represents a human
CC secreted protein of the invention.
XX
XX Sequence 291 AA;
SQ
Query Match      53.2%; Score 1303; DB 22; Length 291;
Best Local Similarity 92.1%; Pred. No. 1.4e-127;
Matches 258; Conservative 4; Mismatches 6; Indels 12; Gaps 3;

QY 27 MDFIQHLGVCCLVALLISVGLLSVAACWFLPSIIAAAASWIIITCVLLCCSKHARCFILLVVF 86
Db 1 MDFIQHLGVCCLVALLISVGLLSVAACWFLPSIIAAAASWIIITCVLLCCSKHARCFILLVVF 60

QY 87 LSCGLREGNALLAAGTGIVILGHVENIFHNPKGLLDGTCNLRKAKSFIHPPLKKYIE 146
Db 61 LSCGLREGNALLAAGTGIVILGHVENIFHNPKGLLDGTCNLRKAKSFIHPPLKKYIE 120

QY 147 AIQWIYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 206
Db 121 AIQWIYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 180

QY 207 LSSLGQKLLAFAGLSVLVLTGTLFMRKFLGPGCWKYENIYITRQVFQDERERHQORPCV 266
Db 181 LSSLGQKLLAFAGLSVLVLTGTLFMRKFLGPGCWKYENIYITRQVFQDERERHQORPCV 240

QY 267 LPLNKEERKYYIIPFTWPTPKERNVLGLFFPILHLCLI 306
Db 241 LPLNKEERK-----NKELEKLSM-IIP-LIYLCL 268

RESULT 13
AAB89984
ID AAB89984 standard; Protein; 148 AA.
XX
AC AAB89984;
XX
XX 24-MAY-2002 (first entry)
XX
XX Human polypeptide SEQ ID NO 2360.
XX
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
XX vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein.
XX
```

OS Homo sapiens.
 XX WO200190304-A2.
 XX 29-NOV-2001.
 XX 18-MAY-2001; 2001WO-US16450.
 XX 19-MAY-2000; 2000US-20551SP.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Birse CE, Rosen CA;
 XX WPI; 2002-122018/16.
 XX N-PSDB; ABL90393.
 DR Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders -
 XX
 XX Claim 11; SEQ ID NO 2360; 2081pp + Sequence Listing; English.
 PS
 XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 148 AA;
 SQ
 Query Match 30.8%; Score 754.5; DB 23; Length 148;
 Best Local Similarity 91.5%; Pred. No. 1.3e-70;
 Matches 140; Conservative 2; Mismatches 4; Indels 7; Gaps 1;
 QY 231 MKRFLGPCGKVENIYITRFQVQFDERHRHQPCVLPINKERRKYVIPTFWPTPKER 290
 Db 1 MKRFLGPCGKVENIYITRFQVQFDERHRHQPCVLPINKERRKYVIPTFWPTPKER 60
 QY 291 KNLGLFFLPILHLCIWLFAAVDYLLYLIFSVSKQFOSLPFGFVHLKHGEKQTQDI 350
 Db 61 KNLGLFFLPILHLCIWLFAAVDYLLYLIFSVSKQFOSLPFGFVHLKHGEKQTQDI 120
 QY 351 IHDSGFNISVFBNCPKPKFLLSETWVPLSVI 383
 Db 121 IHDSGFNISVFBNCPKPKFLLSETWVPLSVI 146
 RESULT 14
 AAY45267
 ID AAY45267 standard; Protein; 485 AA.
 XX
 XX AAY45267;
 XX 23-NOV-1999 (first entry)
 DT Human secreted protein encoded from gene 11.
 XX Secreted protein; human; gene therapy; diagnosis; treatment; cancer;
 KW

protein therapy; tumor; neurodegenerative disorder; blood disorder; AIDS;
 developmental abnormality; leukemia; immune system; autoimmune disease;
 hepatic disease; renal disease; inflammation; allergy; schizophrenia;
 Alzheimer's disease; cognitive disorder; arthritis; infection; psoriasis;
 transplant rejection; diabetes; asthma; sepsis; acne; metabolic disorder;
 cardiovascular disorder; food additive; preservative.
 XX Homo sapiens.
 OS WO9946289-A1.
 XX 16-SEP-1999.
 XX 11-MAR-1999; 99WO-US05721.
 XX 12-MAR-1998; 98US-0077686.
 PR 12-MAR-1998; 98US-0077687.
 PR 12-MAR-1998; 98US-0077696.
 PR 12-MAR-1998; 98US-0077714.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Ruben SM, Ferrie AM, Rosen CA, Florence C, Young PE, Yu G, Ni J;
 PI WPI; 1999-551363/46.
 XX N-PSDB; AAZ27243.
 DR
 XX New isolated human genes, useful for diagnosis and treatment of, e.g.
 PT cancers -
 PT
 PS Claim 1b; 179-180; 306pp; English.
 XX
 XX This invention describes novel human genes and the secreted proteins
 CC they encode. The polynucleotides and their corresponding secreted
 CC polypeptides are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also pathological conditions
 CC can be diagnosed by determining the amount of the new polypeptides in a
 CC sample or by determining the presence of mutations in the new
 CC polynucleotides. Specific uses are described for each of the
 CC polynucleotides of the invention, based on which tissues they are most
 CC highly expressed in, and include developing products for the diagnosis or
 CC treatment of cancer, tumors, neurodegenerative disorders, developmental
 CC abnormalities, blood disorders, leukemias, diseases of the immune system,
 CC autoimmune diseases, hepatic and renal disease, inflammation, allergies,
 CC Alzheimer's and cognitive disorders, schizophrenia, arthritis,
 CC infections, AIDS, transplant rejection, diabetes, asthma, sepsis, acne,
 CC psoriasis, cardiovascular disorders, and metabolic disorders. The
 CC polypeptides or polynucleotides can also be used as food additives or
 CC preservatives. The polypeptides are also useful for identifying their
 CC binding partners. AAY45257-Y45289 represent the human secreted proteins
 CC described in the method of the invention.
 XX
 XX Sequence 485 AA;
 SQ
 Query Match 7.0%; Score 171; DB 20; Length 485;
 Best Local Similarity 20.6%; Pred. No. 8.7e-09;
 Matches 94; Conservative 63; Mismatches 167; Indels 132; Gaps 18;
 QY 35 VCCLVALISVGLLSVAACWFLPSIIAAASWITCVLLCCSKHARCIFILVFCGREG 94
 Db 5 VCGLLVFLSLGL-----VPPV-----RCLFALSVPITLGMEOG 36
 QY 95 RNALIAAGTIVILGHVENIFHNFKGLLDGTCNLRKSPSIHFFLKKYIEALQWYGL 154
 Db 37 RRLLSYSTATLCAIAVVPNVLANVGAAGQVLRCTEGS-----LESLNTHQLHAA 88
 QY 155 ATPLSVFDLLVSNWQTLAVSLFSPSHVLEAQLNDSKGEVLVSILYQMATTEVLVSLG-Q 212
 Db 89 SRALGPTGQAGSRGLT-----FEAQDNGS-----AFYLMHTVTQVLEDFSGLE 133
 QY 213 KLIAFAGLSILVILGTGLFMKRFILGPCGW-----KYENIYITRFVQ--FDERHRQ 261
 Db 134 SLARAAALGTQRVVTGTLFMLGLLVESAWYLHCYLTDLRFDNIVATQTLQRLAQATHL 193

QY 262 QRP-----CVLPNKEERKYYIIPTFWPTPKERNKLGFLFPLILHLICWVFAAV 313
DB 194 LAPPPTWLLQAQURLSOEELLSCLL-----RLGLLAL-LIVATAVAV---AT 237
QY 314 DYLLYRIIFSVMKQFQSLPGFVHLKLGKQGTODIIDSFSNISVFEP---NCIKPK 369
DB 238 DHVAFLLAQATVDWAQKLPTVPI-----TLTVKYDVAYTVLGPFLFNOLAPES 287
QY 370 KFL---LSEW-----VPLSV--ILLILVMLGSLLSILMQLKILV 404
DB 288 PFLSVHSSYQWELRLTSARCPLLPARRPRAAPLXAGGLQALLAGSTVLLEGVARRLXAI 347
QY 405 SASFPVSVERKRIQVILHAKILKRSKQ-----PLGE 435
DB 348 AASFPTAQEARIRHLHARLQRHRDRXGQQLPLGD 383

RESULT 15
ABB65609
ID ABB65609 standard; Protein; 684 AA.
AC ABB65609;
XX
XX
DT 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 23619.
DE Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
KW
XX Drosophila melanogaster.
OS
XX
XX WO200171042-A2.
FN
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX
XX WPI; 2001-656860/75.
DR N-FSDB; ABL09712.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 23619; 21pp + Sequence Listing; English.
PS
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 684 AA;

Query Match 6.5%; Score 158.5; DB 22; Length 684;
Best Local Similarity 17.1%; Pred. No. 2.9e-07;
Matches 91; Conservative 83; Mismatches 157; Indels 201; Gaps 17;

QY 66 IITCVLLCCSKHARCFILAVFLSCGLREGRNALIAAGTGIVILGHVENIFHNFKGLDGM 125
DB 74 LVITGTFMLSLPVRVILLIFVALVCKSGRTYLRAVAFIISGPIANLVENAGEVARVF 133
QY 126 TC-----NLRKSFSSIHFPPLKLYIEAIQWI----- 151
DB 134 VCTTVLTYNLSKTRPDLMAKPTNTLKHMRGDVEEIRHTFYELQEVLDLKYAVENSIE 193
QY 152 ---YG-----LATPLSV----- 160
DB 194 DEKYGDKNTPKYERWGRETSRMNVSEIGNGGKELPTPAQVORFORNMNRCKHQLRSG 253
QY 161 -----FDDLVS-----WNQTLAV-----SLFSPSHVL----- 182
DB 254 HRACLEVFRNGYRKCTTNFPSPMIARCAICWPYRVDIICELDLFGNPKICDPSAVVPQFG 313
QY 183 -----EAQLNDSKGEVLVLYQMATTTTEVLSSLG----- 211
DB 314 ETVVELLKAEEKLFDNSSQIV-VNVEIKDEQFAKSQLKAERTGQAFKEDFERQKRIFNK 372
QY 212 -----OKLAFAGLSLVLLGTGLFMKRFLGPGCWKYENIYITQFVQFDERERHQRP 266
DB 373 VMGILQKILCLPMLRMVYVSINYYVK-YLNDV--EPDNFYITKYFKHVDQRRKEQRI 429
QY 267 LPLNKEERRKYVIIPTFWPTPKERNKLGFLFPLILHLICWVFAAVDYLLYRLFVS 326
DB 430 LPLRTYKSKYIDVD-----H-----IFSRTHHSTTIVRKRSK 463
QY 327 -QFQSLPGFEVHLKLHGEKQGT---QDIIDSSFNI-----SVFEPNCIPKPKFLLSET 376
DB 464 IVYQDQGEHEVRFNISGVGQMARLUTTWLH--NFNIHEKVSTLSNKECLPNNAHVL 521
QY 377 WYPLSVILLILVMLGSLLSILMQLKILVGSASFYPSVERKRIQVILHAKILK 428
DB 522 YVQLILLVLIIVLIYQSTTFELMRVVICSFYFKKEQRIILYLNRLLENR 573

Search completed: June 21, 2003, 13:23:40
Job time : 74 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2003, 13:22:27 ; Search time 27 Seconds
(without alignments)
512.177 Million cell updates/sec

Title: US-09-713-098-2

Perfect score: 2448

Sequence: 1 NGIWTGTDIFLSLWEIYVS.....LPVLKMRKQMDVASADKS 470

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA*
1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A-COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B-COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCJUS-COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93	3.8	337	2	US-08-467-559B-2
2	92	3.8	447	4	US-09-370-253-6
3	91	3.7	1498	2	US-08-404-531B-29
4	91	3.7	1498	3	US-08-476-900A-29
5	91	3.7	1498	3	US-08-488-546A-29
6	91	3.7	1582	2	US-08-404-531B-9
7	91	3.7	1582	3	US-08-476-900A-9
8	91	3.7	1582	3	US-08-488-546A-9
9	91	3.7	1582	4	US-08-726-320-5
10	91	3.7	1582	4	US-09-208-716-5
11	91	3.7	2132	4	US-08-669-656A-6
12	90.5	3.7	489	4	US-09-518-914-4
13	89	3.6	496	4	US-09-134-001C-3001
14	89	3.6	554	4	US-08-180-371-2
15	89	3.6	870	4	US-09-134-001C-4959
16	89	3.6	2336	4	US-09-268-163-10
17	88.5	3.6	2337	1	US-08-455-543A-48
18	88.5	3.6	2237	2	US-08-223-305C-48
19	88.5	3.6	2237	4	US-09-268-163-8
20	88.5	3.6	2296	2	US-08-286-819A-27
21	88.5	3.6	2296	3	US-08-980-357-27
22	88.5	3.6	2337	3	US-08-713-118-2
23	88.5	3.6	2337	4	US-09-452-007-2
24	88.5	3.6	2339	1	US-08-455-543A-47
25	88.5	3.6	2339	2	US-08-223-305C-47
26	88.5	3.6	2339	4	US-09-268-163-6
27	88.5	3.6	2343	4	US-09-268-163-4

28	87.5	3.6	447	4	US-09-370-253-10	Sequence 10, Appl
29	87.5	3.6	823	1	US-07-745-206A-15	Sequence 15, Appl
30	87.5	3.6	823	2	US-08-311-363-15	Sequence 15, Appl
31	87.5	3.6	1754	1	US-07-745-206A-13	Sequence 13, Appl
32	87.5	3.6	1754	2	US-08-311-363-13	Sequence 13, Appl
33	87.5	3.6	1835	3	US-08-836-325-15	Sequence 15, Appl
34	87	3.6	394	2	US-08-742-440A-8	Sequence 8, Appl
35	87	3.6	453	1	US-08-570-157-7	Sequence 7, Appl
36	87	3.6	453	4	US-09-076-510-7	Sequence 7, Appl
37	87	3.6	1498	2	US-08-404-531B-28	Sequence 28, Appl
38	87	3.6	1498	3	US-08-476-900A-28	Sequence 28, Appl
39	87	3.6	1498	3	US-08-488-546A-28	Sequence 28, Appl
40	87	3.6	1581	2	US-08-404-531B-6	Sequence 6, Appl
41	87	3.6	1581	3	US-08-476-900A-6	Sequence 6, Appl
42	87	3.6	1581	3	US-08-488-546A-6	Sequence 6, Appl
43	87	3.6	1581	4	US-08-726-320-4	Sequence 4, Appl
44	87	3.6	1581	4	US-09-208-716-4	Sequence 4, Appl
45	86.5	3.5	219	6	5310729-4	Patent No. 5310729

ALIGNMENTS

RESULT 1

US-08-467-559B-2
; Sequence 2, Application US/08467559B
; Patent No. 5928890
; GENERAL INFORMATION:
; APPLICANT: LI, YI
; TITLE OF INVENTION: HUMAN AMINE RECEPTOR
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN AND FOX, P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,559B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0840000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 337 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-467-559B-2

Query Match 3.8%; Score 93; DB 2; Length 337;
Best Local Similarity 21.4%; Pred. No. 0.1;
Matches 83; Conservative 55; Mismatches 137; Indels 112; Gaps 18;
QY 83 LLVFLSCGLREGNALLAAGTGVILGHVENIFHNFKGLDGMTCNLRAKSF--SIHFP- 139
Db 35 LVIVITC-----AAGMLIIVLG---NVFVF-----AVSYFKALHTPT 69
QY 140 --LLKKYIEAIIQWYGLATPLSV-----FDDLVSNNQITLAVSLF---SPSHVLEAQ 185
Db 70 NEFLLLSLALADMFLGLVLPLSTIRSVESCWFGDFLCRLHLYDLTLFCUTSIFHLCFIS 129

QY 186 LNDKGEVLSVLYQMATTTEVLSLQKLLAFAGLSVLVLTGLFMRFLGPCGKMYENI 245
Db 130 IDHCAICDPLLYPSKFTVRV-----ALRYILAGWGV-----PAA--YTSL 168
QY 246 YITRFQVQDERERHQPCV-----LPNKEERRKVVIIPTWPTPKERKNLGLFLPL 301
Db 169 FLYTDVETRLSQWLEEMPCVSCQLLNK-----FWGW-----LNFPLFPVPL 213
QY 302 IHLICWLVFAADYLLYLIFSVSQFOSLPGFEVHLKLHGKQGTQDIIHDSSENFV 361
Db 214 IMISLYVKIFVATQAQIITLSK---SLAGAAKHERKAAKTLG----- 255
QY 362 BENCIPKPKFLSETWVPLSVLLVMGLSSILMQLKILVSASFYSPVERKRIQYLH 421
Db 256 ---IWGIIYL--CWLPTIDTMDVSLHFTPLV-PDIFWAFYNSACNPIIYVPS 308
QY 422 AKLLKRSKQPLGE-----VKERLSLY 443
Db 309 YOWFRKALKLTLKQKVFSPQTRVDLY 335

RESULT 2

US-09-370-253-6

; Sequence 6, Application US/09370253

; Patent No. 6165792

; GENERAL INFORMATION:

; APPLICANT: Allen, Stephen M.

; APPLICANT: Sakai, Hajime

; APPLICANT: Thorpe, Catherine J.

; TITLE OF INVENTION: Amino Acid Transporters

; FILE REFERENCE: BB-1200

; CURRENT APPLICATION NUMBER: US/09/370,253

; EARLIER FILING DATE: 1999-08-09

; EARLIER APPLICATION NUMBER: 60/097,222

; EARLIER FILING DATE: August 20, 1998

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 6

; LENGTH: 447

; TYPE: PRT

; ORGANISM: Oryza sativa

; US-09-370-253-6

Query Match 3.8%; Score 92; DB 4; Length 447;
Best Local Similarity 21.6%; Pred. No. 0.2;
Matches 55; Conservative 34; Mismatches 88; Indels 78; Gaps 12;

QY 15 WEIYVSPRSGMDFIOHGLGVCLVALISVGLLS-----VAACWFLPSIIIAAASWITCV 70
Db 26 WLPITSSRNAKWWYSAFH---NVTAMVGAGVLSLPYAMSELGWGFGIAVLILSMIITLY 81
QY 71 LLC-----CSKHARFILLVFLSCGLREG-----RNALIAAGTGI 105
Db 82 TLWQVMEHEMVPGRRFORYHELGOHA-----FGEKGLGLMTIWPQQLVVEGVNI 131
QY 106 VIL---GHVENIFHNFKLLDGMTC-NLRKASF-----SIHPELLK-KYIEAIQWLYGLA 155
Db 132 VTMVTGGSKLKFHDV--LCEHGCKNKHLYTFIMIFASVHFVLSQLPNFNSISGVSLAA 189
QY 156 TPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV-----LSS 209
Db 190 AVMSLSYSTIANGASV-----DKGKADVVDYHLRATTSTGKVGFFSA 232
QY 210 LQKLLAFAGLSVL 224
Db 233 LGDVAFAVAGHNVL 247

RESULT 3

US-08-404-531B-29

; Sequence 29, Application US/08404531B

; Patent No. 5863724

; GENERAL INFORMATION:

; APPLICANT: Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson, Pamela

; APPLICANT: Thomas, Gilbert Cote, and Robert Gagel

; TITLE OF INVENTION: Sequence Encoding Mammalian Sulfonylurea Receptor

; Patent No. 5863724

; NUMBER OF SEQUENCES: 49

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &

; ADDRESSEE: No. 5863724ris

; STREET: One Liberty Place 46th. Floor

; CITY: Philadelphia

; STATE: PA

; COUNTRY: USA

; ZIP: 19103

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/404,531B

; FILING DATE: 15-MAR-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Beardell, Lori Y.

; REGISTRATION NUMBER: 34,293

; REFERENCE/DOCKET NUMBER: BYLR-0003

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 215-568-3100

; TELEFAX: 215-568-3439

; INFORMATION FOR SEQ ID NO: 29:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1498 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-404-531B-29

Query Match 3.7%; Score 91; DB 2; Length 1498;
Best Local Similarity 18.6%; Pred. No. 1.6;
Matches 96; Conservative 75; Mismatches 168; Indels 176; Gaps 25;

QY 65 WIITCVLLCCSKHARCFILLV-----FLSCGLREGNRLIAAGTGVILGHVENI--FHN 117
Db 75 WILTFILL-----FVLVCEIAGILSDGVTESRHLHLYMPAGMAFMAITSVVYHN 126
QY 118 FKGLDGMTCNLRKASFHFPLKLYIEAIQWYGLATP---LSVFDLVSNQOT--- 170
Db 127 IE-----TSNFPKLLIAL-LIYWTAFITTKIKFVKFYDHAIGFSQLRFC 170
QY 171 ---LAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEVLSLQKLLAFAGLSVLVLT 227
Db 171 LTGLLVILYGLMLLVENVIRVRYIFFTKTPREVKKPEDQLDGLVRFQ--PFVNLLSKGT 229
QY 228 GLFMKRFGLG-----PCGWKYENIYITRFQVQFDERERHQ-QRP----- 264
Db 230 YWMNAPFKTAHKPIDLRAIKLPIAMRALTYN-QELCVAFDAQAKDTQSPQARAIW 288
QY 265 -----CVL-----PLNKEERRKVVIIPTWPTPKER 290
Db 289 RALCHAFGRRLILSFTIRLADLLGFPAGLPCIFGIVDHLGKENH-----VFQP---KT 338
QY 291 KNIGLFF-----LPILHLCIWLFAVDYLLYRLIFSVSKQPSQLPGFVHL 338
Db 339 QFLGVYFVSSQEFGLGNAYYVAVLLFLAL-----LQRTTFLQSYV---AIBTGI 385
QY 339 KLHGEKQGT--QDI IHDSSFNISV-----FEPNCIPKPKFLLSETW-VPLSV 382
Db 386 NLNGAIQTLYNKIMEMSTNSLNGEWTAQICNLVAIDTNQLMWFFFLCNLTMPVQI 445
QY 383 I---LLILVMLGLLSSILMQLKILVSASFYSPVERKRIQYLHAKLLKRSKQPLGEVRR 439
Db 446 IGVVILLYYILGVSALIGAIVILLAP-----VOYFVATKLSAQRTTLEHSNER 495

QY 440 L-----SLYTKIHFWLPV-----LKMIRKKOM 462
DB 496 LKQTNEMLRGMKLLKYAWESIFCSRVEVTRKEM 530

RESULT 4
US-08-476-900A-29
; Sequence 29, Application US/08476900A
; Patent No. 6031150
; GENERAL INFORMATION:
; APPLICANT: Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson
; TITLE OF INVENTION: Sequence Encoding Mammalian Sulfonylurea Receptor
; Patent No. 6031150
; TITLE OF INVENTION: and Method of Detecting Persistent Hyperinsulinemic Hypoglycemia
; TITLE OF INVENTION: Infancy
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: No. 6031150ris
; STREET: One Liberty Place 46th. Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,900A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; REFERENCE/DOCKET NUMBER: BYLR-0027
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1498 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-476-900A-29

Query Match 3.7%; Score 91; DB 3; Length 1498;
Best Local Similarity 18.6%; Pred. No. 1.6;
Matches 96; Conservative 75; Mismatches 168; Indels 176; Gaps 25;

QY 65 WIITCVLLCCSHKARCFILLV-----FLSCGLREGNRLAIAAGTGIVILGHVENI--FHN 117
DB 75 WILTFILL-----FVLVCEIAEGILSDGVTSRHLHLYPAGMAFMAAITSVVVYHN 126

QY 118 PKGLDGMTCNLRKASFHSIHFPPLKKYIEATQWIYGLATP-----LSVFDDLSVWNOT--- 170
DB 127 IE-----TSNPFKLLIAL-LIYWTIAFTTKIKFKFYDHAIGFSQLRFC 170

QY 171 ---LAVLSPSPSHVLEAGLNDKSGVLSVLYQMATTTEVLSLGGKLLAFAGLSLVLLGT 227
DB 171 LTGLLVILYGMILLVENVNIRVRIYFPKTPREVKPPEDQLDQVRFLO-PFVNLSLKG 229

QY 228 GLFMKRFGL-----PCGWKENIYITQFVQDERERHQ-QRP----- 264
DB 230 YWMNAFIKTAHKPIDLRATAKPIAMRALNY-QRLCVAFDAQKDXQTQSPQARAIM 288

QY 265 -----CVL-----PLNKEERYKVIIPTFWTPK 290
DB 289 RALCHAFGRRLILSTFRILADLLGAFGLFCIFGIVDHLGKNH-----VFQP---KT 338

QY 291 KNLGLFF-----LPILHLICIWVLPAAVDYLLVRLIFSVSKQFSLPGRVHL 338
DB 339 QFLGVTFVSSQBEFLGNAYVLAVLLFAL-----LIQTEFLQASYYV-----AIEGTI 385

QY 339 KLHGKQGT--QDIHDSFNISSV-----FEPNCIPKPKFLILSETW-VPLSV 382
DB 386 NLRGALQTKIYNKIMHMSTSNLSMGEMTAGOICNLVAIDTNQLMWFPLCPNLMTMPVQI 445

QY 383 I---LLILVMLGSLSSILMQLKILVSASPYSPVERKRIQVYLHAKLLKKRSKQPLGEVRR 439
DB 446 IGVGVILYYILGVSALIGAIVILLAP-----VQYFVATKLSQAQRTTLEHSNER 495

QY 440 L-----SLYTKIHFWLPV-----LKMIRKKOM 462
DB 496 LKQTNEMLRGMKLLKYAWESIFCSRVEVTRKEM 530

RESULT 5
US-08-488-546A-29
; Sequence 29, Application US/08488546A
; Patent No. 6054313
; GENERAL INFORMATION:
; APPLICANT: Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson, Pamela
; APPLICANT: Thomas, Gilbert Cole, and Robert Gagel
; TITLE OF INVENTION: Sequence Encoding Mammalian Sulfonylurea Receptor
; Patent No. 6054313
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: No. 6054313ris
; STREET: One Liberty Place 46th. Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,546A
; FILING DATE: 07-JUNE-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/404,531
; FILING DATE: 15-MARCH-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; REFERENCE/DOCKET NUMBER: BYLR-0026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1498 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-488-546A-29

Query Match 3.7%; Score 91; DB 3; Length 1498;
Best Local Similarity 18.6%; Pred. No. 1.6;
Matches 96; Conservative 75; Mismatches 168; Indels 176; Gaps 25;

QY 65 WIITCVLLCCSHKARCFILLV-----FLSCGLREGNRLAIAAGTGIVILGHVENI--FHN 117
DB 75 WILTFILL-----FVLVCEIAEGILSDGVTSRHLHLYPAGMAFMAAITSVVVYHN 126

QY 118 PKGLDGMTCNLRKASFHSIHFPPLKKYIEATQWIYGLATP-----LSVFDDLSVWNOT--- 170

Db 127 IE-----TSNPKLLIAL-LIYWTIAFITKTKFYDHAIGFSQLRFC 170
QY 171 ---LAVSPSPSHVLEAQLNDSKGEVLSVLYQMATTTEVLSSLGOKLLAFAGLSVLVLTG 227
Db 171 LTGLVILYGLMILLVEVNVIRVRYIFFTKPREVKPPEDLDGLGVRFLO-PFVNLLSKGT 229
QY 228 GLEMKRFLG-----PCGMKYENIYITROFVQDERERHQ-ORP-----264
Db 230 YWMNAFIKTAHKPIDLRAIAKLPAMRALTNY-ORLCVAFDAQARKDTSQOGARAIW 288
QY 265 -----CVL-----PLNKEERRKYVIIPTFTWPTPKER 290
Db 289 RALCHAFGRRLILSSTFRILADLLGFAGPLCTIFGIVDHLGKENH-----VFQP---KT 338
QY 291 KNLGLFF-----LPILIHLCIWLFAAVDYLLYRLFVSQFOSLPGEVHL 338
Db 339 QFLGVYFVSSQBFGLNAYVLAVLLFLAL-----LLQRTFLQASYYV---AIETGI 385
QY 339 KLHGEKQGT--ODIHDSSFNISV-----FEPNCIPKPKFLLSETW-VPLSV 382
Db 386 NLRGAIQTKIYNKIMHMTSNLSMGEMTAGQICNLVAIDTNQLMWFFFLCPNLMTMPVQI 445
QY 440 L-----SLYTKIHFWLPV-----LKMIRKKOM 462
Db 496 LKQTNEMLRGMKLLKLYAWESIFCSRVEVTRKEM 530

RESULT 6

US-08-404-531B-9
; Sequence 9, Application US/08404531B
; Patent No. 5863724
; GENERAL INFORMATION:
; APPLICANT: Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson, Pamela
; APPLICANT: Thomas, Gilbert Cote, and Robert Gagel
; TITLE OF INVENTION: Sequence Encoding Mammalian Sulfonylurea Receptor
; Patent No. 5863724
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: No. 5863724ris
; STREET: One Liberty Place 46th. Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/404,531B
; FILING DATE: 15-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardsell, Lori Y.
; REGISTRATION NUMBER: 34,293
; REFERENCE/DOCKET NUMBER: BYLR-0003
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1582 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-404-531B-9

Query Match 3.7%; Score 91; DB 2; Length 1582;
Best Local Similarity 18.6%; Pred. No. 1.8; 168; Indels 176; Gaps 25;
Matches 96; Conservative 75; Mismatches 75; Indels 176; Gaps 25;
QY 65 WIITCVLLCCSHKARCFILLV-----FLSCGLREGNALIAAGTGVILGHVENI--FHN 117
Db 75 WLTIFILL-----FVLVCEIAEGILSDGVTSRHLHLYMPAGNAFMAAITSVVYHN 126
QY 118 FKGLDGMTCNLRKASFSIHFPLLKKYIEAIOIYGLATP-----LSVFDLVSNNQT--- 170
Db 127 IE-----TSNPKLLIAL-LIYWTIAFITKTKFYDHAIGFSQLRFC 170
QY 171 ---LAVSPSPSHVLEAQLNDSKGEVLSVLYQMATTTEVLSSLGOKLLAFAGLSVLVLTG 227
Db 171 LTGLVILYGLMILLVEVNVIRVRYIFFTKPREVKPPEDLDGLGVRFLO-PFVNLLSKGT 229
QY 228 GLEMKRFLG-----PCGMKYENIYITROFVQDERERHQ-ORP-----264
Db 230 YWMNAFIKTAHKPIDLRAIAKLPAMRALTNY-ORLCVAFDAQARKDTSQOGARAIW 288
QY 265 -----CVL-----PLNKEERRKYVIIPTFTWPTPKER 290
Db 289 RALCHAFGRRLILSSTFRILADLLGFAGPLCTIFGIVDHLGKENH-----VFQP---KT 338
QY 291 KNLGLFF-----LPILIHLCIWLFAAVDYLLYRLFVSQFOSLPGEVHL 338
Db 339 QFLGVYFVSSQBFGLNAYVLAVLLFLAL-----LLQRTFLQASYYV---AIETGI 385
QY 339 KLHGEKQGT--ODIHDSSFNISV-----FEPNCIPKPKFLLSETW-VPLSV 382
Db 386 NLRGAIQTKIYNKIMHMTSNLSMGEMTAGQICNLVAIDTNQLMWFFFLCPNLMTMPVQI 445
QY 383 I---LLILVMLGLSSILMQLKILVSASFYPSVERKRIQYLHAKLLKRSKQPLGEVKRR 439
Db 446 IGVVILYYILGVSAIGAAVILLAP-----VQYFVATKLSQAQRTTLEHSNER 495
QY 440 L-----SLYTKIHFWLPV-----LKMIRKKOM 462
Db 496 LKQTNEMLRGMKLLKLYAWESIFCSRVEVTRKEM 530

RESULT 7

US-08-476-900A-9
; Sequence 9, Application US/08476900A
; Patent No. 6031150
; GENERAL INFORMATION:
; APPLICANT: Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson
; TITLE OF INVENTION: Sequence Encoding Mammalian Sulfonylurea Receptor
; Patent No. 6031150
; TITLE OF INVENTION: and Method of Detecting Persistent Hyperinsulinemic Hypoglycemia
; TITLE OF INVENTION: Infancy
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: No. 6031150ris
; STREET: One Liberty Place 46th. Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,900A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardsell, Lori Y.
; REGISTRATION NUMBER: 34,293
; REFERENCE/DOCKET NUMBER: BYLR-0027

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,546A
; FILING DATE: 07-JUNE-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/404,531
; FILING DATE: 15-MARCH-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; REFERENCE/DOCKET NUMBER: BYLR-0026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1582 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-488-546A-9

      3.7%; Score 91; DB 3; Length 1582;
      Best Local Similarity 18.6%; Pred. No. 1.8;
      Matches 96; Conservative 75; Mismatches 168; Indels 176; Gaps 25

Qy    65 WIITCVLLCCSKHARCIFILV-----FLSCGLREGRNALIAAGTGIVILGHVENI--PHN 117
Db    75 WILTFILL-----FVLVCETAEGLSDGVTESRHLHYMPAGMAFMATITSVVYHN 126
Qy   118 PKGLLDGMCNLRAKSFIIHFPLKKYIEALOWNYGLATP----LSVFDDLVSNNQT--- 170
Db   127 IE-----TSNFPKLIAL-LIYWTLAFITTKIKFYKYDHAIGFSQLRFC 170
Qy   171 --LAVSLFSPSHVLEAQLNDSKEVLSVLVQMATTTEVLSSLGOKLLAFAGLSLVLLGT 227
Db   171 LTGLLVLYGMALLVEVNVIRRYIIFKTPREVKPPEDLODLGVRFLO-PFVNLLSGKT 229
Qy   228 GLFMKRLPG-----PCGWKYENIYITQFVQDFDERRHQ-QRP----- 264
Db   230 YWMWNAFIKTAHKKPIDLRAIKLPIDAMRALTNV-QRLCVAFDAQAARKDTQSOGARAIW 288
Qy   265 -----CVL-----PLNKBERKXVIIPTFWPTPKER 290
Db   289 RALCHAFGRRLILSSTPRILADLLGFAGPLCIFGIVDHKGKNH-----VFQP---KT 338
Qy   291 KNIGLFP-----LPILIHLCIWLVFAADVLLYRLIPSVSKQPSQLPGFEVHL 338
Db   339 QFLGVFEVSSQEFNGNAVYLAVLLFTAL-----LLORTFLOAQSYV-----ALETGI 385
Qy   339 KLHGCKQGT--QDIHHDSSEFNISV-----FEPNCIPKPFLLSETW-VPLSV 382
Db   386 NLRGAIO TKYNKIMHMSTNSLWGEMTAGQICNLVAIDTNQLMWFFELCNLTWMPVOI 445
Qy   383 I----LTI LVMGLSSLIMOLKILVSAFYSFVERKEIQYLHAKLLKKRSKQPLGEVKRR 439
Db   446 IVGVILLYILGYVALIGA AVIILLAP-----VQYFVATKLSAQORTTLEHSNER 495
Qy   440 L-----SLYLTKHFWLPV-----LKMTKKQM 462
Db   496 LKOTNMLRGMKLLKLYAWESIFCSRVVEVTNRREM 530

```

```

339 QFGLGVFVSSQEPFLGNAIVLAVLLEFLAL-----LLQKTFPLQASYYV-----ALETGI 381
Qy 339 KLHGEKQGT--QDIHHDSFNISV-----FEPNCIPKPKFLISETW-VPLSV 382
Db 386 NLRGAIQTKYINKIMHMSTENLSNGEMTAGOICNLVAIDTNQLMWFFFLCPNLMTMPVQI 445
Qy 383 I-----LLLTLMGLLSLTLMOLKIIVSASFPSYVERKRIQYLHAKLLKKRSKQPLGEVKR 439
Db 446 IGVVILYYILGVSAIGAAVILLAP-----VQYFVATKLSQAORTTLEHSNER 495
Qy 440 L-----SLYLTKHFWLPV-----LKMIRKQM 462
Db 496 LKQTNEMLRGMKLLKLYAWESI FCSRVVEVTRRKEM 530

RESULT 9
US-08-726-320-5
; Sequence 5, Application US/08726320
; Patent No. 6171815
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Bandman, Olga

```

APPLICANT: Coleman, Roger
TITLE OF INVENTION: NOVEL HUMAN SULFONYLUREA RECEPTOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726,320
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: AF-0001 US
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1582 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 784874
US-08-726-320-5

Query Match 3.7%; Score 91; DB 4; Length 1582;
Best Local Similarity 18.6%; Pred. No. 1.8; Indels 176; Gaps 25;
Matches 96; Conservative 75; Mismatches 168; Indels 176; Gaps 25;
QY 65 WIITCVLLCCSKHRCFILLV-----FLSCGLREGNALLAAGTGIVILGHVENI--FHN 117
DB 75 WILTFILL-----FVLVCEIAEGILSDGVTSRHLHLYMPAGMAFMAAITSVVYHN 126
QY 118 FKGLLDGMTCNLRKSFHSHPFLKXKYEIAQIYGLATP-----LSVPDDLVSNNQT--- 170
DB 127 IE-----TSNPFKLLIAL-LIYWTLAFITKTKFVKFYDHAIGFSQLRFC 170
QY 171 ---LAVSLFSPSHVLEAQLNDSKGEVLSVLYQYATTTVEVLSLQKLLAPAGLSVLIGT 227
DB 171 LTGLLVILYGMILLVEVNVIRVRYFFKTPREVKKPEDQLDGVRFLO-PFVNLLSKGT 229
QY 228 GLFMKRFGLG-----PCGWKENYIITRQVQDFDERRHQ-QRP----- 264
DB 230 YWMAFIKTAHKFIDLRALKLPIAMRALTYN-ORLCVADQARKDTQSPQGARATW 288
QY 265 -----CVL-----PLNKERRKYVIPTFWPTPKR 290
DB 289 RALCHAFGRRLISSTFRILADLLGAPGLCIFGIVDLGKENH-----VFOP---KT 338
QY 291 KNLGLFF-----LPILIHLCIWLFAANDYLYRLIFSVKSFQSLPGFEVHL 338
DB 339 QPLGVYFVSSQBFNGAYVLAIVLLFLAL-----LLQRTFLOASYV----ALETGI 385
QY 339 KLHGEKQGT--QDIHDSFNLSV-----FEPNCIPKPKFILLSETW-VPLSV 382
DB 386 NURGAIQTKYINKMHMTSNLSMGEMTAGOICNLVAIDTNQLMWFELCPNLWTPVQI 445
QY 383 I---LLILVMLGILLSSILMQLKILVSASFYPSVERKRIQVYLHAKLLKRSKQPLGEVKRR 439

DB 446 IVGVILYYILGVSAIGAAVILLAP-----VQYFVATKLSQAORTLHRSNER 495
QY 440 L-----SLYLTIKIHFPLPV-----LKMIRKKQM 462
DB 496 LKQTNEMLRGMKLLKLYAWESIFCSRVETRRKEM 530
RESULT 10
US-09-208-716-5
Sequence 5, Application US/09208716
Patent No. 6235880
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
APPLICANT: Bandman, Olga
APPLICANT: Coleman, Roger
TITLE OF INVENTION: NOVEL HUMAN SULFONYLUREA RECEPTOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/208,716
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/726,320
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: AF-0001 US
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1582 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 784874
US-09-208-716-5

Query Match 3.7%; Score 91; DB 4; Length 1582;
Best Local Similarity 18.6%; Pred. No. 1.8; Indels 176; Gaps 25;
Matches 96; Conservative 75; Mismatches 168; Indels 176; Gaps 25;
QY 65 WIITCVLLCCSKHRCFILLV-----FLSCGLREGNALLAAGTGIVILGHVENI--FHN 117
DB 75 WILTFILL-----FVLVCEIAEGILSDGVTSRHLHLYMPAGMAFMAAITSVVYHN 126
QY 118 FKGLLDGMTCNLRKSFHSHPFLKXKYEIAQIYGLATP-----LSVPDDLVSNNQT--- 170
DB 127 IE-----TSNPFKLLIAL-LIYWTLAFITKTKFVKFYDHAIGFSQLRFC 170
QY 171 ---LAVSLFSPSHVLEAQLNDSKGEVLSVLYQYATTTVEVLSLQKLLAPAGLSVLIGT 227
DB 171 LTGLLVILYGMILLVEVNVIRVRYFFKTPREVKKPEDQLDGVRFLO-PFVNLLSKGT 229
QY 228 GLFMKRFGLG-----PCGWKENYIITRQVQDFDERRHQ-QRP----- 264

Db 230 YWMNAFIKTAHKPIDLRAIAKPIAMRALTNV-QLRCVAFDAQARKDQTSPQGARAIW 288
Qy 265 -----CUL-----PLNKEERKRYIIPFTFWTPKER 290
Db 289 RALCHAFGRRLISLSTFRILADLLGFAGPLCIFGIVDHLGKHN-----VFQP-----KT 338
Qy 291 KNLGLFF-----LPLIHLICWLFPAADVLLYRLIFSVSKQFQSLPGFEVHL 338
Db 339 QFLGVYFVSSQBFELGNVAVLVLFLAL-----LLQRTFLQASYV-----AIETGI 385
Qy 339 KLHGKQGT--QDIHDSFNISV-----PEPNCIPKPKFILLSETW-VPLSV 382
Db 386 NLRGAIQTKYKNKIMHSTSNLSMGEMTAGQICNLVAIDTNQLMWFELCPNLWTMPQVI 445
Qy 383 I-----LLILVNLGLLSILMQLKILVASFPYSVERKRIQYLHAKLLKRSKQPLGEVKRR 439
Db 446 IGVVILLYILGVSALIGAAVILLAP-----VQYFVATKLSQAQRTTLSHSNER 495
Qy 440 L-----SLYTKIHFWLPV-----LKMWRKKOM 462
Db 496 LKQTNEMLRGMKLLKLYAWESIFCSRVEVTRRKEM 530

RESULT 11

US-08-669-656A-6
; Sequence 6, Application US/08669656A
; Patent No. 6451554
; GENERAL INFORMATION:
; APPLICANT: Wood, John N.
; APPLICANT: Akopian, Armen N.
; TITLE OF INVENTION: Ion Channel
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESS: ZENECA Pharmaceuticals
; STREET: 1800 Concord Pike, P.O. Box 15437
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: USA
; ZIP: 19850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/669,656A
; FILING DATE: 24-JUN-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Hohenschutz, Liza D.
; REGISTRATION NUMBER: 33,712
; REFERENCE/DOCKET NUMBER: PHM.70086
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 886-7466
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2132 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-669-656A-6

Query Match 3.7%; Score 91; DB 4; Length 2132;

Best Local Similarity 19.3%; Pred. No. 2.8; Mismatches 121; Indels 148; Gaps 23;
Matches 81; Conservative 69;

Qy 56 PSIIAAASWIITCVLLCCSHARC-----FILLVFLSCGLREGRNALIAAGTG 104
Db 633 PCLISFAQYLI---WECCPKWRKFKMALFELVTDPPFAELITLC-----IVNVTV 680
Qy 105 IVLGHVENIFHNFKLLDGMTCNLRAKS--FSHFPLKKY-IEAIIQWYGLATPLSVF 161
Db 681 FWAMEHYV-----MTDAFDAMLQAGNIVFTVFTMEMAFKIIAFDPYFYFQKKWNIF 732

Qy 162 DDLVSNQTLAYSLFSPSHVLEAQLNDSKEVLSVLVYQMATTTEVLSSLOKLLAFAG-- 219
Db 733 DCVI-----VTVSLL-----ELSASKGSLSVLRSRL--ALDTTQKSFLSAGYL 776
Qy 220 -----LSVLVLLGTGLFMRKRFGLPCGKWKENYITRQFVQFDERERHQORPCVPLPN 270
Db 777 NEPPFRAQRAMSVSVISWV-----IEELESKUKCPCLLSF- 813
Qy 271 KEERKRYIIPFWP--TPKERK-NLGLF-----FLPLIHLICIW--LFAAVDY----- 315
Db 814 ---AQYLI---WECCPKWRKFKMALFELVTDPPFAELTITLCIVVNTVFMAMEHYPMWD 866
Qy 316 -----LLYRLIFSVSKQFQSLPGFEVH-----LKHGKQK 346
Db 867 AFDAMLQAGNIVFTVFTMEMAFKII-AFDPPYFYFQKKWNIFDCVIVTVSLLLSASKG 925
Qy 347 TQDIHDSFNISVPEPNCIPKPKFILLSETWVPLSVLLIL-VMLGILLSILMQLKLV 404
Db 926 SLVLSRL-DRVFK-----LAKSWPTLNTLIKIGNSVGALNLTFLILAIIV 973

RESULT 12

US-09-518-914-4
; Sequence 4, Application US/09518914
; Patent No. 6413731
; GENERAL INFORMATION:
; APPLICANT: Borowsky, Beth E.
; APPLICANT: Ogozalek, Kristine L.
; APPLICANT: Lakhiani, Parul P.
; APPLICANT: Adham, Nika
; TITLE OF INVENTION: DNA ENCODING SNORF36a AND SNORF36b RECEPTORS
; FILE REFERENCE: 59138-A/JPW
; CURRENT APPLICATION NUMBER: US/09/518,914
; CURRENT FILING DATE: 2000-03-03
; EARLIER APPLICATION NUMBER: US 09/303,593
; EARLIER FILING DATE: 1999-05-03
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-518-914-4

Query Match 3.7%; Score 90.5; DB 4; Length 489;

Best Local Similarity 19.6%; Pred. No. 0.34; Mismatches 84; Conservative 59; Mismatches 129; Indels 157; Gaps 20;

Qy 20 SPRSPGWMDFIQHLGVCCILVALISVGLLSVAACWFLPSIIAA-----AASWIITCVLLCC 74
Db 22 TPAPSWDSSQS-----SISLGRLPSPAPTGTWAAW-VPLPTVDV 65
Qy 75 SKHARCFILLVFLSCGLREGRNALIAAGTGVILGHVENIFHNFKLLDGMTCNLRAKSF 134
Db 66 PDHAHYTLGTVILLVGL-----TG--MLGNLTVIYTCRAVLRGVTVMMQSR- 111
Qy 135 SIHFPLKKYIEAIIQWYGLATP-----LSVFDLVSNNQTLAVLSLSPSHVLEAQLN 187
Db 112 -----LRTPANMFIINLAVSDFMSFTQ---APVFTSLYKQWLF 149
Qy 188 DSKG-----EVLVLYQMATTTEVLSSIGOKL-----LAFAGL-----SLVLGTGL 229
Db 150 GETGCFEYAFGALFGISSMITLTAI--ALDRYIVTRPLATFGVAKRRAAVFLGVWL 207
Qy 230 F-----MKRFLG-----PGWKYENIYITRQFVQFDERERHQORPCVPLPNKEE 273
Db 208 YALAWSLPPFFGWSAVVPEGLTSCSWYMSF-----TPAV----- 243
Qy 274 RKYVVIIPFTWTPKERKNLGLFFLPLIHLICWLFPAADVLLYRLIFSVSKQFQSLPG 333
Db 244 -RAYTMLLCCF-----VFPLPLLIIVCYI-----FIFRAIRETGRALQTFGA 285

Search completed: June 21, 2003, 13:27:03
Job time : 29 secs

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2002, 08:08:19 ; Search time 33.44 Seconds
(without alignments)
1070.635 Million cell updates/sec

Title: US-09-713-098-2
Perfect score: 470
Sequence: 1 MGWTSGTDFLSWEIYVS.....LPVLKMKRKQMDMASADKS 470

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : PIR_68.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	1.7	177	B82221	conserved hypoteth
2	8	1.7	272	E64182	cys2 protein - Hae
3	8	1.7	338	T35517	probable secreted
4	8	1.7	408	S17793	cyclin B - common
5	8	1.7	461	D96835	probable cyclin, 4
6	8	1.7	1285	B72420	hypothetical prote
7	7	1.5	64	D69412	hypothetical prote
8	7	1.5	67	A25188	H+-transporting AT
9	7	1.5	78	C81239	hypothetical prote
10	7	1.5	82	JU0350	9.4K protein - chl
11	7	1.5	122	E96802	unknown protein [i
12	7	1.5	126	S74493	thioredoxin M-2 -
13	7	1.5	126	B46538	Ig heavy chain, me
14	7	1.5	137	S52361	superoxide dismuta
15	7	1.5	150	D69100	probable protein d
16	7	1.5	153	T19054	hypothetical prote
17	7	1.5	158	H83824	hypothetical prote
18	7	1.5	160	D70189	conserved hypoteth
19	7	1.5	162	C81879	probable membrane
20	7	1.5	168	S69480	succinate dehydrog
21	7	1.5	180	E71337	hypothetical prote
22	7	1.5	182	T15753	hypothetical prote
23	7	1.5	188	A75382	hypothetical prote
24	7	1.5	190	A37753	DNA repair protein
25	7	1.5	190	T19841	hypothetical prote
26	7	1.5	201	C53331	guanylate cyclase
27	7	1.5	227	T42160	probable kfr prote
28	7	1.5	235	B34342	IgE Fc receptor be
29	7	1.5	250	T18862	hypothetical prote

30	7	1.5	251	2	T09980	H+-transporting AT
31	7	1.5	253	2	T19775	hypothetical prote
32	7	1.5	262	2	T26366	hypothetical prote
33	7	1.5	264	2	I38136	chymotrypsin-like
34	7	1.5	268	2	E83460	probable transcrip
35	7	1.5	278	2	S75883	hypothetical prote
36	7	1.5	279	1	C29336	ubiquinol--cytochr
37	7	1.5	280	1	C25405	ubiquinol--cytochr
38	7	1.5	285	2	S63260	hypothetical prote
39	7	1.5	303	2	E81348	binding-protein de
40	7	1.5	306	2	A85814	probable transmemb
41	7	1.5	306	2	JS0266	membrane protein y
42	7	1.5	318	2	A86638	conserved hypoteth
43	7	1.5	322	2	C82991	probable aromatic
44	7	1.5	325	2	T12348	NADH dehydrogenase
45	7	1.5	335	2	H70545	probable fabH prot
46	7	1.5	340	2	H72325	rod shape-determin
47	7	1.5	345	2	T12347	NADH dehydrogenase
48	7	1.5	345	2	T12346	NADH dehydrogenase
49	7	1.5	352	2	F70983	probable serine pr
50	7	1.5	361	2	S47170	hypothetical prote

ALIGNMENTS

RESULT 1

B82221
conserved hypothetical protein VC1274 [imported] - Vibrio cholerae (strain Ni6961 ser
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: B82221
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833

A:Accession: B82221

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-177 <HEI>

A:Cross-references: GB:AE004206; GB:AE003852; NID:g9555749; PIDN:AAF94433.1; GSPDB:GN
A:Experimental source: serogroup O1; strain Ni6961; biotype El Tor

C:Genetics:

A:Gene: VC1274

A:Map position: 1

Query Match 1.7%; Score 8; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 390 LGLLSSIL 397

|||||||

Db 18 LGLLSSIL 25

RESULT 2

E64182

cys2 protein - Haemophilus influenzae (strain Rd KW20)

C:Species: Haemophilus influenzae

C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 20-Aug-1999

C:Accession: E64182

R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A:Reference number: A64000; MUID:95350630

A:Accession: E64182

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-272 <TIGR>
A:Cross-references: GB:032790; GB:IL42023; NID:g1574654; PIDN:AAC22757.1; PID:g1574656; T
C:Genetics:
A:Function:
A:Description: probably involved in sulfate transport
A:Pathway: cysteine biosynthesis
C:Superfamily: cysZ protein
C:Keywords: cysteine biosynthesis; membrane protein; sulfate transport

Query Match 1.7%; Score 8; DB 2; Length 272;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 LSVILLIL 387
|||||||
Db 70 LSVILLIL 77

RESULT 3
T35517
probable secreted protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C:Accession: T35517
A:Submitted to the EMBL Data Library, March 1999
A:Reference number: 221581
A:Accession: T35517
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-338 <SEE>
A:Cross-references: EMBL:AL049497; PIDN:CAB39862.1; GSPDB:GN00070; SCOEDB:SC6G10.08c
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC6G10.08c

Query Match 1.7%; Score 8; DB 2; Length 338;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 VGLLSVAA 51
|||||||
Db 20 VGLLSVAA 27

RESULT 4
S17793
cyclin B - common limpet
C:Species: Patella vulgata (common limpet)
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 16-Jul-1999
C:Accession: S17793; S16136
R:van Loon, A.E.; Collas, P.; Goedemans, H.J.; Neant, I.; Dalbon, P.; Guerrier, P.
EMBO J. 10, 3343-3349, 1991
A:Title: The role of cyclins in the maturation of Patella vulgata oocytes.
A:Reference number: S17792; MUID:92007785
A:Accession: S17793
A:Molecule type: mRNA
A:Residues: 1-408 <LOO>
A:Cross-references: EMBL:X58358; NID:g10954; PIDN:CAA41255.1; PID:g10955
C:Superfamily: cyclin
C:Keywords: cell cycle control

Query Match 1.7%; Score 8; DB 2; Length 408;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 PSIIAAAA 63
|||||||

Db 323 PSIIAAAA 330

RESULT 5
D96835
probable cyclin, 42214-44381 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: D96835
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alcon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: D96835
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-461 <STO>
A:Cross-references: GB:AE005173; NID:g6751715; PIDN:AAF27697.1; GSPDB:GN00141
C:Genetics:
A:Gene: F5I6.12
A:Map position: 1
C:Superfamily: cyclin

Query Match 1.7%; Score 8; DB 2; Length 461;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 FLPSIIAA 61
|||||||
Db 373 FLPSIIAA 380

RESULT 6
B72420
hypothetical protein TM0088 - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: B72420
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hic
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
A:Reference number: A72200; MUID:99287316
A:Accession: B72420
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1285 <ARN>
A:Cross-references: GB:AE001695; GB:AE000512; NID:g4980569; PIDN:AAD35182.1; PID:g498
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0088

Query Match 1.7%; Score 8; DB 2; Length 1285;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 383 ILLILVML 390
|||||||
Db 7 ILLILVML 14

RESULT 7
D69412

hypothetical protein AF1301 - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: D69412
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
A:Reference number: A69250; MUID:98049343
A:Accession: D69412
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-64 <KLE>
A:Cross-references: GB:AE001014; GB:AE000782; NID:g2689337; PIDN:AAB89953.1; PID:g264928

Query Match 1.5%; Score 7; DB 2; Length 64;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 LISVGLL 47
|||||||
DB 44 LISVGLL 50

RESULT 8

A25188 H+-transporting ATP synthase (EC 3.6.1.34) protein 8 - Chinese hamster mitochondrion

N:Alternate names: A6L protein
C:Species: mitochondrion Crictetus griseus (Chinese hamster)

C:Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 07-Dec-1999
C:Accession: A25188

R:Breen, G.A.M.; Miller, D.L.; Holmans, P.L.; Welch, G.
J. Biol. Chem. 261, 11680-11685, 1986

A:Title: Mitochondrial DNA of two independent oligomycin-resistant Chinese hamster ovary
A:Reference number: A25188; MUID:86304297
A:Accession: A25188

A:Molecule type: DNA
A:Residues: 1-67 <BRE>

A:Cross-references: GB:M14311; NID:g336719; PIDN:AAA68615.1; PID:g336720

C:Genetics:

A:Genome: mitochondrion

A:Genetic code: SGC1

C:Superfamily: H+-transporting ATP synthase protein 8

C:Keywords: ATP biosynthesis; hydrolase; membrane-associated complex; mitochondrion; oxi

Query Match 1.5%; Score 7; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 396 ILMQLKI 402
|||||||
DB 22 ILMQLKI 28

RESULT 9

C81239 hypothetical protein NMB0091 [imported] - Neisseria meningitidis (strain MC58 serogroup

C:Species: Neisseria meningitidis

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C:Accession: C81239

R:Rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve

A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755

A:Accession: C81239

Query Match 1.5%; Score 7; DB 2; Length 122;

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-78 <TE>

A:Cross-references: GB:AE002368; GB:AE002098; NID:g7225303; PIDN:AAF40553.1; PID:g722

A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NMB0091

Query Match 1.5%; Score 7; DB 2; Length 78;

Best Local Similarity 100.0%; Pred. No. 31;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 322 FSVSKQF 328
|||||||
DB 51 FSVSKQF 57

RESULT 10

JU0350

9.4K protein - Chlamydomophila psittaci phage Chp1

C:Species: Chlamydomophila psittaci phage Chp1, Chlamydia psittaci phage Chp1

A:Note: host Chlamydia psittaci

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jun-2000

C:Accession: JU0350

R:Storey, C.C.; Lusher, M.; Richmond, S.J.

J. Gen. Virol. 70, 3381-3390, 1989

A:Title: Analysis of the complete nucleotide sequence of Chp1, a phage which infects

A:Reference number: JU0345; MUID:90111716

A:Molecule type: DNA

A:Accession: JU0350

A:Residues: 1-82 <STO>

A:Cross-references: GB:D00624; NID:g217761; PIDN:BAA00508.1; PID:g217767

C:Superfamily: Chlamydia psittaci phage Chp1 9.4K protein

Query Match 1.5%; Score 7; DB 1; Length 82;

Best Local Similarity 100.0%; Pred. No. 33;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 384 LLILVML 390
|||||||
DB 38 LLILVML 44

RESULT 11

E96802

unknown protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: E96802

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: E96802

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-122 <STO>

A:Cross-references: GB:AE005173; NID:g11079487; PIDN:AAG29199.1; GSPDB:GN00141

C:Genetics:

A:Gene: F2p24.6

A:Map position: 1

Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 ALTAAGT 103
|||||
Db 78 ALTAAGT 84

RESULT 12

S74493

thioredoxin M-2 - Synchocystis sp. (strain PCC 6803)

N:Alternate names: protein sll1057

C:Species: Synchocystis sp.

A:Variety: PCC 6803

C:Accession: S74493

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synchocystis

S.

A:Reference number: S74322; MUID:97061201

A:Accession: S74493

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-126 <KAN>

A:Cross-references: EMBL:D90899; GB:AB001339; NID:gi1651650; PIDN:BAAL6645.1; PID:gi165171

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Gene: trxm-2

C:Superfamily: thioredoxin; thioredoxin homology

Query Match

Best Local Similarity 1.5%; Score 7; DB 2; Length 126;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 AVSLFSP 178
|||||
Db 120 AVSLFSP 126

RESULT 13

B46538

Ig heavy chain, membrane-bound form - Atlantic cod

C:Species: Gadus morhua (Atlantic cod)

C:Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000

A:Accession: B46538; S15479

R:Benigien, E.; Leanderson, T.; Pilstrom, L.

Eur. J. Immunol. 21, 3027-3033, 1991

A:Title: Immunoglobulin heavy chain cDNA from the teleost Atlantic cod (Gadus morhua L.)

A:Reference number: A46538; MUID:92083930

A:Accession: B46538

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-126 <BEN>

A:Cross-references: EMBL:X58871; NID:g63914; PIDN:CAA41681.1; PID:g63915

A:Note: sequence extracted from NCBI backbone (NCBIP:72613)

C:Keywords: alternative splicing

Query Match

Best Local Similarity 1.5%; Score 7; DB 2; Length 126;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 GEVLSVL 197
|||||
Db 25 GEVLSVL 31

RESULT 14

S52361

superoxide dismutase (EC 1.15.1.1) (Mn) - Mycobacterium chelonae (NCTC 0946, isolate Frl

N:Alternate names: manganese superoxide dismutase
C:Species: Mycobacterium chelonae
A:Variety: strain NCTC 0946, isolate Friedmann type
C:Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 05-Mar-1999
C:Accession: S52361
R:Bull, T.T.

Submitted to the EMBL Data Library, January 1995

A:Description: Rapid identification of mycobacteria from AIDS patients by capillary e

A:Reference number: S52360

A:Accession: S52361

A:Molecule type: DNA

A:Residues: 1-137 <BUL>

A:Cross-references: EMBL:Z48216

A:Experimental source: NCTC 0946, isolate Friedmann type

C:Genetics:

A:Gene: Mn-SOD

C:Function:

A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxy

C:Superfamily: superoxide dismutase (Mn)

C:Keywords: metalloprotein; oxidoreductase

Query Match 1.5%; Score 7; DB 2; Length 137;

Best Local Similarity 100.0%; Pred. No. 51;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 SLGQKLL 215
|||||
Db 105 SLGQKLL 111

RESULT 15

D69100

probable protein disulfide-isomerase (EC 5.3.4.1) MTH1745 [similarity] - Methanobacte

C:Species: Methanobacterium thermoautotrophicum

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Sep-2000

C:Accession: D69100

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.

Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan,

ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: f

A:Reference number: A69000; MUID:98037514

A:Accession: D69100

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-150 <MTH>

A:Cross-references: GB:AE000930; GB:AE000666; NID:g2622872; PIDN:AAB6215.1; PID:g262

A:Experimental source: strain Delta H

C:Genetics:

A:Gene: MTH1745

C:Keywords: intramolecular oxidoreductase; isomerase

Query Match

Best Local Similarity 1.5%; Score 7; DB 2; Length 150;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 AGLSLVL 224
|||||
Db 19 AGLSLVL 25

RESULT 16

T19054

hypothetical protein C07E3.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000

C:Accession: T19054

R:Matthews, P.

submitted to the EMBL Data Library, June 1995

A:Reference number: T19056

A:Accession: T19054

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-153 <WIL>
A:Cross-references: EMBL:Z49908; PIDN:CAA90098.1; GSPDB:GN000020; CESP:C07E3.9
A:Experimental source: clone C07E3
C:Genetics:
A:Gene: CESP:C07E3.9
A:Map position: 2
A:Introns: 42/3; 81/3; 107/1
C:Superfamily: phospholipase A2

Query Match 1.5%; Score 7; DB 2; Length 153;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 FILLVFL 87
|||||
Db 3 FILLVFL 9

RESULT 17
H83824
hypothetical protein BH1400 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000
C:Accession: H83824
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20263314
A:Accession: H83824
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-158 <STO>
A:Cross-references: GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BAB05119.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH1400

Query Match 1.5%; Score 7; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 295 LFFLPIL 301
|||||
Db 110 LFFLPIL 116

RESULT 18
D70189
conserved hypothetical integral membrane protein BB0717 - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: D70189
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943
A:Accession: D70189
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-160 <KLE>
A:Cross-references: GB:AE001171; GB:AE000783; NID:g2688640; PIDN:AAC67057.1; PID:g268864
A:Experimental source: strain B31

Query Match 1.5%; Score 7; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 58;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 391 GLLSSIL 397
|||||
Db 49 GLLSSIL 55

RESULT 19
C81879
probable membrane protein NM1128 [imported] - Neisseria meningitidis (strain Z2491 s
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: C81879
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo
; Holroyd, S.; Jagers, K.; Leather, S.; Moulé, S.; Mungall, K.; Quail, M.A.; Rajandre
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491
A:Reference number: A81775; MUID:2022556
A:Accession: C81879
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-162 <PAR>
A:Cross-references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84389.1; PID:g7737
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NM1128

Query Match 1.5%; Score 7; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 ALIAAGT 103
|||||
Db 11 ALIAAGT 17

RESULT 20
S68480
succinate dehydrogenase homolog YLR164w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein I9632.1
C:Species: Saccharomyces cerevisiae
C:Date: 20-Jul-1996 #sequence_revision 23-Aug-1996 #text_change 05-Nov-1999
C:Accession: S68480
R:Vaudin, M.
submitted to the EMBL Data Library, July 1996
A:Description: The sequence of S. cerevisiae cosmid 9362.
A:Reference number: S68471
A:Accession: S68480
A:Molecule type: DNA
A:Residues: 1-168 <VAU>
A:Cross-references: EMBL:U51921; NID:g1234843; GSPDB:GN00012; MIPS:YLR1
C:Genetics:
A:Gene: MIPS:YLR164w
A:Map position: 12R
C:Keywords: transmembrane protein
F:56-83/Domain: transmembrane #status predicted <TM1>
F:124-140/Domain: transmembrane #status predicted <TM2>

Query Match 1.5%; Score 7; DB 2; Length 168;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 LSLVLLG 226
|||||
Db 93 LSLVLLG 99

RESULT 21
E71337
hypothetical protein TP0335 - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)

C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
C:Accession: E71337

R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwinn
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.

A:Title: Complete genome sequence of *Treponema pallidum*, the syphilis spirochete.

A:Reference number: A71250; MUID:98332770

A:Accession: E71337

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-180 <COL>

A:Cross-references: GB:AE001213; GB:AE000520; NID:93322606; PIDN:AA65328.1; PID:9332261

A:Experimental source: strain Nichols

C:Genetics:

A:Gene: TP0335

C:Superfamily: syphilis spirochete hypothetical protein TP0335

Query Match 1.5%; Score 7; DB 2; Length 180;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 191 GEVLSVL 197

Db 113 GEVLSVL 119

RESULT 22

T15753

hypothetical protein C33F10.11 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T15753

R:Chissoe, S.

submitted to the EMBL Data Library, February 1996

A:Description: The sequence of *C. elegans* cosmid C33F10.

A:Reference number: Z18397

A:Accession: T15753

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-182 <CHI>

A:Cross-references: EMBL:U49830; NID:gl203933; PID:gl203943; PIDN:AAA93401.1; CESP:C33F1

C:Genetics:

A:Gene: CESP:C33F10.11

A:Introns: 138/3

Query Match 1.5%; Score 7; DB 2; Length 182;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 380 LSVILLI 386

Db 18 LSVILLI 24

RESULT 23

A75382

hypothetical protein - *Deinococcus radiodurans* (strain R1)

C:Species: *Deinococcus radiodurans*

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: A75382

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

S.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.

A:Reference number: A75250; MUID:20036896

A:Accession: A75382

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-188 <WHI>

A:Cross-references: GB:AE001999; GB:AE000513; NID:96459316; PIDN:AAF1123.1; PID:9645

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR1559

A:Map position: 1

Query Match 1.5%; Score 7; DB 2; Length 188;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 220 LSVLLG 226

Db 124 LSVLLG 130

RESULT 24

A37753

DNA repair protein sms - *Salmonella typhimurium* (fragment)

N:Alternate names: DNA repair protein rada

C:Species: *Salmonella typhimurium*

C>Date: 19-Apr-1991 #sequence_revision 30-Apr-1999 #text_change 30-Apr-1999

C:Accession: A37753

R:Foster, J.W.; Park, Y.K.; Penfound, T.; Fenger, T.; Spector, M.P.

J. Bacteriol. 172, 4187-4196, 1990

A:Title: Regulation of NAD metabolism in *Salmonella typhimurium*: molecular sequence a

A:Reference number: A37753; MUID:90330519

A:Accession: A37753

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 'MSWAS', 1-190 <FOS>

A:Cross-references: GB:M33722

C:Genetics:

A:Gene: sms; rada

C:Function:

A:Description: imparts resistance to the alkylating agent methylmethane sulfonate

C:Superfamily: DNA repair protein sms

C:Keywords: ATP; DNA binding; DNA repair; zinc finger

Query Match 1.5%; Score 7; DB 2; Length 190;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 158 LSVFDDL 164

Db 184 LSVFDDL 190

RESULT 25

T19841

hypothetical protein C39B10.4 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000

C:Accession: T19841

R:White, S.

submitted to the EMBL Data Library, July 1995

A:Reference number: Z19186

A:Accession: T19841

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-190 <WIL>

A:Cross-references: EMBL:250027; PIDN:CAA90333.1; GSPDB:GN00028; CESP:C39B10.4

A:Experimental source: clone C39B10

C:Genetics:

A:Gene: CESP:C39B10.4

A:Map position: X

A:Introns: 68/1; 109/1; 160/3

C:Superfamily: *Caenorhabditis elegans* hypothetical protein C39B10.4

Query Match 1.5%; Score 7; DB 2; Length 190;
Best Local Similarity 100.0%; Pred. No. 67;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 KRRLSLY 443
|||||||

Db 48 KRRLSLY 54

RESULT 26
C55331
guanylate cyclase activator 1A - human
N:Alternate names: retinal guanylate cyclase-activating protein
C:Species: Homo sapiens (man)
C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 20-Sep-1999
C:Accession: C55331; A55331
R:Subbaraya, I.; Ruiz, C.C.; Helekar, B.S.; Zhao, X.; Gorczyca, W.A.; Pettenati, M.J.; R
J. Biol. Chem. 269, 31080-31089, 1994
A:Title: Molecular characterization of human and mouse photoreceptor guanylate cyclase-a
A:Reference number: A55331; MUID:95074147
A:Accession: C55331
A:Molecule type: DNA
A:Residues: 1-201 <SUB1>
A:Cross-references: GB:L36861; NID:G623404; PIDN:AAA60542.1; PID:G623405
A>Note: authors translated the codon AAT for residue 185 as His
A:Accession: A55331
A:Molecule type: mRNA
A:Residues: 1-201 <SUB2>
A:Cross-references: GB:L36859; NID:G623402; PIDN:AAA60541.1; PID:G623403
C:Genetics:
A:Gene: GDB:GUCA1A; GUCA1; GUCA; GCAP; GCAP1
A:Cross-references: GDB:132137; OMIM:600364
A:Map position: 6p21.1-6p21.1
A:Introns: 67/3; 117/3; 149/1
C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
F:14-46/Domain: calmodulin repeat homology <EF1>
F:51-83/Domain: calmodulin repeat homology <EF2>
F:87-119/Domain: calmodulin repeat homology <EF3>
F:131-163/Domain: calmodulin repeat homology <EF4>

Query Match 1.5%; Score 7; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 AGLSLVL 224
|||||||

Db 78 AGLSLVL 84

RESULT 27
T42160
probable kfr protein - Escherichia coli plasmid p0157
C:Species: Escherichia coli
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Nov-2000
C:Accession: T42160; T00236
R:Burland, V.; Shao, Y.; Perna, N.T.; Plunkett, G.; Sofia, H.J.; Blattner, F.R.
Nucleic Acids Res. 26, 4196-4204, 1998
A:Title: The complete DNA sequence and analysis of the large virulence plasmid of Esche
A:Reference number: Z22068; MUID:98391744
A:Accession: T42160
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-227 <BUR>
A:Cross-references: EMBL:AF074613; PIDN:AAC70128.1
A:Experimental source: strain EDL933; serotype O157:H7
R:Makino, K.; Ishii, K.; Yasunaga, T.; Hattori, M.; Yokoyama, K.; Yatsudo, H.C.; Kubota,
S.; Shinagawa, H.
DNA Res. 5, 1-9, 1998
A:Title: Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrha
A:Reference number: Z14127; MUID:98290540
A:Accession: T00236
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA

A:Residues: 7-227 <MAK>
A:Cross-references: EMBL:AB011549; PIDN:BAA31783.1
A:Experimental source: strain EHEC O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: KfrAs
A:Genome: plasmid p0157
A>Note: L7060
C:Superfamily: Escherichia coli probable kfr protein

Query Match 1.5%; Score 7; DB 2; Length 227;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 ALIAAGT 103
|||||||

Db 46 ALIAAGT 52

RESULT 28
B34342
Ige Fc receptor beta chain - mouse
C:Species: Mus musculus (house mouse)
C>Date: 15-Jun-1990 #sequence_revision 15-Jun-1990 #text_change 05-Nov-1999
C:Accession: B34342
R:Ra, C.; Jouvin, M.H.E.; Kinet, J.P.
J. Biol. Chem. 264, 15323-15327, 1989
A:Title: Complete structure of the mouse mast cell receptor for IgE (Fc-epsilon-RI) a
A:Reference number: A34342; MUID:89359361
A:Accession: B34342
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-235 <RAC>
A:Cross-references: GB:J05019; NID:G193238; PIDN:AAA37601.1; PID:G309225
C:Keywords: immunoglobulin receptor; transmembrane protein

Query Match 1.5%; Score 7; DB 2; Length 235;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 IAAGTGI 105
|||||||

Db 131 IAAGTGI 137

RESULT 29
T18862
hypothetical protein C02D4.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C:Accession: T18862
R:Baynes, C.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19033
A:Accession: T18862
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-250 <WIL>
A:Cross-references: EMBL:Z81031; PIDN:CAB02717.1; GSPDB:GN000028; CESP:C02D4.1
A:Experimental source: clone C02D4
C:Genetics:
A:Gene: CESP:C02D4.1
A:Map position: X
A:Introns: 71/1; 98/3; 134/3; 182/3
C:Superfamily: Caenorhabditis elegans hypothetical protein C02D4.1

Query Match 1.5%; Score 7; DB 2; Length 250;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 QTLAVSL 175

Db 184 QTLAVSL 190
|||||||

RESULT 30
T09980
H+-transporting ATP synthase (EC 3.6.1.34) chain A - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 24-Nov-1999
C:Accession: T09980
R:Robison, K.
submitted to the EMBL Data Library, September 1994
A:reference number: Z16911
A:Accession: T09980
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-251 <ROB>
A:Cross-references: EMBL:U15186; NID:g699323; PID:g699349
C:Superfamily: H+-transporting ATP synthase protein 6
C:Keywords: ATP biosynthesis; hydrolase; membrane-associated complex

Query Match 1.5%; Score 7; DB 2; Length 251;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 203 TTEVLSS 209
|||||||
Db 119 TTEVLSS 125

Search completed: January 31, 2002, 08:12:01
Job time: 222 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2002, 08:11:24 ; Search time 21.47 Seconds
(without alignments)
802.630 Million cell updates/sec

Title: US-09-713-098-2
Perfect score: 470
Sequence: 1 MGWTSCTGDIPLSLWEIYVS.....LPVLKMKRKQMDASADKS 470

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 100059 seqs, 36664827 residues

Word size : 0

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	1.7	96	1 NULM_MXGXL	Q9g2x0 myxine glut
2	8	1.7	272	1 CYSZ_HAEIN	P45039 haemophilus
3	8	1.7	408	1 CG2B_PATVU	P24862 patella vul
4	8	1.7	790	1 ATSY_SYN7	P37385 synechococc
5	7	1.5	64	1 YD01_ARCFU	O28968 archaeoglob
6	7	1.5	67	1 ATP8_CRIGR	P14414 cricetus
7	7	1.5	67	1 ATP8_MICPE	P24949 microtus pe
8	7	1.5	82	1 Y9KD_BFCHP	P19191 bacterioph
9	7	1.5	138	1 SODM_MYCCH	P53644 mycobacteri
10	7	1.5	180	1 Y335_TREPA	O83355 treponema p
11	7	1.5	200	1 GCAL_HUMAN	P43080 homo sapien
12	7	1.5	224	1 OXO2_HORVU	P45851 hordeum vul
13	7	1.5	235	1 FCBE_MOUSE	P20490 mus musculu
14	7	1.5	251	1 ATP6_MYCLE	P45829 mycobacteri
15	7	1.5	253	1 PSAT_CAEEL	Q95005 caenorhabdi
16	7	1.5	264	1 CTRL_HUMAN	P40313 homo sapien
17	7	1.5	279	1 CY1_RHOCA	P08501 rhodobacter
18	7	1.5	285	1 YN26_YEAST	P53830 saccharomyc
19	7	1.5	298	1 RADA_SALTY	P24517 salmonella
20	7	1.5	299	1 ICIA_AERSA	P70773 aeromonas s
21	7	1.5	306	1 YEDA_ECOLI	P09185 escherichia
22	7	1.5	307	1 EFTS_BARQU	Q9xcm5 bartonella
23	7	1.5	314	1 CBSB_SULSO	P58030 sulfolobus
24	7	1.5	318	1 Y234_AQUAE	O66423 aquifex aeo
25	7	1.5	335	1 FABL_MYCTU	O06399 mycobacteri
26	7	1.5	336	1 CMST_CRIGR	O08520 cricetus
27	7	1.5	336	1 CMST_MOUSE	O61420 mus musculu
28	7	1.5	375	1 YHCM_ECOLI	P46442 escherichia
29	7	1.5	388	1 HCAT_HAEIN	P44629 haemophilus
30	7	1.5	432	1 YF10_MYCTU	P71789 mycobacteri
31	7	1.5	449	1 DHE2_CLOSY	P24295 clostridium
32	7	1.5	452	1 PTCC_ECOLI	P17334 escherichia
33	7	1.5	460	1 RADA_ECOLI	P24554 escherichia

34 7 1.5 466 1 CRD1_ANSAN P33110 anser anser
35 7 1.5 475 1 YLEO_ECOLI P31474 escherichia
36 7 1.5 481 1 YAAJ_HAEIN P44555 haemophilus
37 7 1.5 482 1 ARCD_PSEAE P18275 pseudomonas
38 7 1.5 540 1 ASD1_NEUCR P78710 neurospora
39 7 1.5 544 1 GPI0_DICDI Q06885 dictyosteli
40 7 1.5 555 1 DNLI_ARCFU O29632 archaeoglob
41 7 1.5 565 1 HEMA_IATKP P03456 influenza a
42 7 1.5 571 1 SECD_MYCLE P38387 mycobacteri
43 7 1.5 596 1 NUOL_CAMJE Q9pma7 campylobact
44 7 1.5 604 1 YFIC_BACSU P54719 bacillus su
45 7 1.5 633 1 IPAA_SHIFL P18010 shigella fl
46 7 1.5 658 1 SOHC_ZYMMO P33990 zymomonas m
47 7 1.5 691 1 COMA_NEIGO P51973 neisseria g
48 7 1.5 846 1 YEG1_MYCTU O53152 mycobacteri
49 7 1.5 869 1 YEG1_MYCLE Q49689 mycobacteri
50 7 1.5 997 1 ATC2_CANFA O46674 c sarcoplas

ALIGNMENTS

RESULT 1

NULM_MXGXL
ID NULM_MXGXL STANDARD; PRT; 96 AA.
AC Q9G2X0: O83922;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 4L (EC 1.6.5.3).
GN MTND4L OR ND4L OR ND4H4L.
OS Myxine glutinosa (Atlantic hagfish).
OG Eukaryotia.
OC Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniiformes;
OC Myxiniidae; Myxiniinae; Myxine.
OX NCBI_TaxID:7769;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98210228; PubMed=9541532;
RA Rasmussen A.S., Janke A., Arnason U.;
RT "The mitochondrial DNA molecule of the hagfish (Myxine glutinosa) and
vertebrate phylogeny.";
RL J. Mol. Evol. 46:382-388(1998).
[2]
SEQUENCE FROM N.A.
RA Delarbre C., Rasmussen A.S., Arnason U., Gachelin G.;
RP "Complete sequence of the mitochondrial DNA of Myxine glutinosa.";
RT Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.
CC -|- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

CC EMBL; Y15191; CAA75490.1; ALT_INIT.
DR EMBL; AJ404477; CAC20657.1; -.
DR InterPro; IPR003214; Mit_NADHub_oxidredctse_4L.
DR InterPro; IPR001133; Oxidored_q2.
DR Pfam; PF00420; oxidored_q2; 1.
DR ProDom; PD000359; Mit_NADHub_oxidredctse_4L; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 96 AA; 10501 MW; 862D807596D6CB26 CRC64;

Query Match 1.7%; Score 8; DB 1; Length 96;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 218 AGLSLVLL 225


```

Db 69 AGLSLVLL 76
|||||
RESULT 2
CYS2_HAEIN STANDARD; PRT; 272 AA.
AC P45039;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CYS2 PROTEIN HOMOLOG.
GN CYS2 OR H11102.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus
OX NCBI_TaxID=727;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=RD / KW20 / ATCC 51907;
RC MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Karlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geohagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL Science 269:496-512(1995).
CC -1- FUNCTION: POSSIBLY INVOLVED IN SULFATE TRANSPORT.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC (POTENTIAL).
CC -1- SIMILARITY: TO E. COLI CYS2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32790; AAC22757.1;
DR TIGR; H11102;
KW Cysteine biosynthesis; Transmembrane; Inner membrane; Transport;
KW Complete proteome.
FT TRANSMEM 29 49 POTENTIAL.
FT TRANSMEM 66 86 POTENTIAL.
FT TRANSMEM 87 107 POTENTIAL.
FT TRANSMEM 148 168 POTENTIAL.
FT TRANSMEM 219 239 POTENTIAL.
SQ SEQUENCE 272 AA; 31075 MW; 3CC4ECB5B5047AA3 CRC64;

Query Match 1.7%; Score 8; DB 1; Length 272;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 LSVILLIL 387
Db 70 LSVILLIL 77
|||||
RESULT 3
CG2B_PATVU STANDARD; PRT; 408 AA.
AC P24862;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)

20-AUG-2001 (Rel. 40, Last annotation update)
G2/MITOTIC-SPECIFIC CYCLIN B.
Patella vulgata (Common limpet).
Eukaryota; Metazoa; Mollusca; Gastropoda; Archaeogastropoda;
Patellacea; Patellidae; Patella.
NCBI_TaxID=6465;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=Oocyte;
RC MEDLINE=92007785; PubMed=1655419;
RA van Loon A.E., Colas P., Goedemans H.J., Neant I., Dalbon P.,
RA Guerrier P.;
RT "The role of cyclins in the maturation of Patella vulgata oocytes.";
RL EMBO J. 10:3343-3349(1991).
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G2/M
CC (MITOSIS) TRANSITION.
CC -1- SUBUNIT: INTERACTS WITH THE CDC2 PROTEIN KINASE TO FORM A
CC SERINE/THREONINE KINASE HOLOENZYME COMPLEX ALSO KNOWN AS
CC MATURATION PROMOTING FACTOR (MPF). THE CYCLIN SUBUNIT IMPARTS
CC SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC -1- DEVELOPMENTAL STAGE: ACCUMULATES STEADILY DURING G2 AND IS
CC ABRUPTLY DESTROYED AT MITOSIS.
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN AB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X58358; CAA41255.1;
DR PIR; S17793; S17793.
DR HSP; P20248; IJST.
DR InterPro; IPR000553; Cyclin.
DR Pfam; PF00134; cyclin; 1.
DR SMART; SM00385; CYCLIN; 2.
DR PROSITE; PS00292; CYCLINS; 1.
KW Cyclin; Cell cycle; Cell division; Mitosis.
SQ SEQUENCE 408 AA; 45896 MW; DEB834HDB273BA71 CRC64;

Query Match 1.7%; Score 8; DB 1; Length 408;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 PSIIAAA 63
Db 323 PSIIAAA 330
|||||
RESULT 4
ATSY_SYNP7 STANDARD; PRT; 790 AA.
AC P37385;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROBABLE COPPER-TRANSPORTING ATPASE SYNA (EC 3.6.3.4).
GN SYNA.
OS Synecococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=1140;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=95023970; PubMed=7937823;
RA Phung L.T., Ajlani G., Haselkorn R.;
RT "P-type ATPase from the cyanobacterium Synecococcus 7942 related to
RT the human Menkes and Wilson disease gene products.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:9651-9654(1994).
CC -1- FUNCTION: INVOLVED IN COPPER TRANSPORT.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + CU(2+)(IN) = ADP + PHOSPHATE +

```

CC CU(2+)(OUT).

CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY

CC (E1-E2 ATPASES). SUBFAMILY IB.

CC -!- SIMILARITY: CONTAINS 1 HEAVY-METAL-ASSOCIATED (HMA) DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; U04356; AAB82020.1; -.

DR InterPro; IPR001757; E1-E2_ATPase.

DR InterPro; IPR001934; HMA.

DR InterPro; IPR001454; Hydrolase.

DR Pfam; PF00122; E1-E2_ATPase; 1.

DR Pfam; PF00403; HMA; 1.

DR Pfam; PF00702; Hydrolase; 1.

DR PROSITE; PS00154; ATPASE_E1_E2; 1.

DR PROSITE; PS01047; HMA; 1.

KW Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding;

KW Copper; Metal-binding.

FT DOMAIN 1 105

FT TRANSMEM 106 127

FT DOMAIN 128 136

FT TRANSMEM 137 156

FT DOMAIN 157 163

FT TRANSMEM 164 184

FT DOMAIN 185 198

FT TRANSMEM 199 219

FT DOMAIN 220 358

FT TRANSMEM 359 381

FT DOMAIN 382 420

FT TRANSMEM 421 438

FT DOMAIN 439 723

FT TRANSMEM 724 743

FT DOMAIN 744 755

FT TRANSMEM 756 774

FT DOMAIN 775 790

FT DOMAIN 790 790

FT MOD_RES 476 476

FT METAL 669 669

FT METAL 673 673

FT METAL 25 25

FT METAL 28 28

SQ SEQUENCE 790 AA; 83694 MW; FA5B7BFEDF4C3793 CRC64;

Query Match 1.7%; Score 8; DB 1; Length 790;

Best Local Similarity 100.0%; Pred. No. 12;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 221 SLVLGTG 228

Db 170 SLVLGTG 177

|||||||

RESULT 5

ID YD01_ARCFU STANDARD; PRT; 64 AA.

AC Q28968;

DT 20-AUG-2001 (Rel. 40, Created)

DT 20-AUG-2001 (Rel. 40, Last sequence update)

DE HYPOTHETICAL PROTEIN AFI301.

GN AFI301.

OS Archaeoglobus fulgidus.

OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;

OC Archaeoglobus.

OX NCBI_TaxID=2234;

RP SEQUENCE FROM N.A.

RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;

RX MEDLINE=98049343; PubMed=9389475;

RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,

RA Kitchum K.A., Dodson R.J., Winn M., Hickey E.K., Peterson J.D.,

RA Reichmann D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,

RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,

RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,

RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,

RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,

RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,

RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,

RA Mason T.N., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,

RA Venter J.C.;

RT "The complete genome sequence of the hyperthermophilic, sulphate-

RT reducing archaeon *Archaeoglobus fulgidus*.";

RL Nature 390:364-370(1997).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF001014; AAB89953.1; -.

DR TIGR; AFI301.

KW Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 30 52

SQ SEQUENCE 64 AA; 7395 MW; 1F5991D80B100A6F CRC64;

Query Match 1.5%; Score 7; DB 1; Length 64;

Best Local Similarity 100.0%; Pred. No. 14;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 LISVGLL 47

Db 44 LISVGLL 50

|||||||

RESULT 6

ID ATP8_CRIGR STANDARD; PRT; 67 AA.

AC P14414;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE ATP SYNTHASE PROTEIN 8 (EC 3.6.1.34) (ATPASE SUBUNIT 8) (A6L).

GN MTATP8 OR ATP8

OS *Cricetulus griseus* (Chinese hamster).

OG Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;

OC *Cricetulus*.

OX NCBI_TaxID=10029;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=86304297; PubMed=3017940;

RA Breen G.A.M., Miller D.L., Holmans P.L., Welch G.;

RT "Mitochondrial DNA of two independent oligomycin-resistant Chinese

RT hamster ovary cell lines contains a single nucleotide change in the

RT ATPase 6 gene.";

RL J. Biol. Chem. 261:11680-11685(1986).

CC -!- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT

CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.

CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND.

CC -!- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M14311; AAA68615.1; --
DR PIR: A25188; A25188.
DR InterPro: IPR001421; ATP-synt_8.
DR InterPro: IPR003238; Mamm_mito_ATPase_8.
DR Pfam: PF00895; ATP-synt_8; 1.
DR ProDom: PD001090; Mamm_mito_ATPase_8; 1.
KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
FT TRANSMEM 8 24 POTENTIAL.
SQ SEQUENCE 67 AA; 7905 MW; E61396FB48137487 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 67;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 396 ILMQLKI 402
|||||
DB 22 ILMQLKI 28

RESULT 7
ID ATP8_MICPE STANDARD; PRT; 67 AA.
AC P24949;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE ATP SYNTHASE PROTEIN 8 (EC 3.6.1.34) (ATPASE SUBUNIT 8) (A61L).
GN MTATP8 OR ATP8
OS Microtus pennsylvanicus (Meadow vole).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Arvicolinae;
OC Microtus.
OX NCBI_TaxID=10058;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92211731; PubMed=1313508;
RA Pumo D.E., Phillips G.J., Barcia M., Millan C.;
RT "Three patterns of mitochondrial DNA nucleotide divergence in the
RT meadow vole, *Microtus pennsylvanicus*."
RL J. Mol. Evol. 34:163-174(1992).
CC -!- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND.
CC -!- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X60285; CAA42826.1; --
DR InterPro: IPR001421; ATP-synt_8.
DR InterPro: IPR003238; Mamm_mito_ATPase_8.
DR Pfam: PF00895; ATP-synt_8; 1.
DR ProDom: PD001090; Mamm_mito_ATPase_8; 1.
KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
FT TRANSMEM 8 24 POTENTIAL.
SQ SEQUENCE 67 AA; 7817 MW; 716105D86A149911 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 67;
Best Local Similarity 100.0%; Pred. No. 15;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 396 ILMQLKI 402
|||||
DB 22 ILMQLKI 28

RESULT 8
ID Y9KD_BPCHP STANDARD; PRT; 82 AA.
AC P19191;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DE HYPOTHETICAL 9.4 KDA PROTEIN (ORF6).
OS Bacteriophage Chpl.
OC Viruses.
OX NCBI_TaxID=12367;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90111716; PubMed=2607341;
RA Storey C.C., Lusher M., Richmond S.J.;
RT "Analysis of the complete nucleotide sequence of Chpl, a phage which
RT infects avian Chlamydia psittaci."
RL J. Gen. Virol. 70:3381-3390(1989).
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: D00624; BAA00508.1; --
DR PIR: J00350; J00350.
KW Hypothetical protein.
SQ SEQUENCE 82 AA; 9426 MW; 9639C09F9229C121 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 82;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 384 LLILVLM 390
|||||
DB 38 LLILVLM 44

RESULT 9
ID SODM_MYCCH STANDARD; PRT; 138 AA.
AC P53644;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE SUPEROXIDE DISMUTASE [MN] (EC 1.15.1.1) (FRAGMENT).
GN SODA OR SOD.
OS Mycobacterium chelonae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1774;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCCT 0946 / FRIEDMANN;
RA Bull T.J., Shanson D.C., Archard L.C.;
RT "Rapid identification of mycobacteria from AIDS patients by capillary
RT electrophoretic profiling of amplified SOD gene."
RL J. Clin. Pathol. Clin. Mol. Pathol. 48:124-132(1995).
CC -!- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
CC -!- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).
CC -!- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE

CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z48216; CAA88249.1; -.
 CC HSSP; P17670; LIDS.
 CC InterPro: IPR001189; SOD_MI.
 CC Pfam: PF00081; sodfe; 1.
 CC ProDom: PD000475; SOD_MI; 1.
 CC PROSITE; PS00088; SOD_MN; PARTIAL.
 CC Oxidoreductase; Manganese.
 CC FT NON_TER 1 1
 CC FT METAL 1 1 MANGANESE (BY SIMILARITY).
 CC FT METAL 49 49 MANGANESE (BY SIMILARITY).
 CC FT METAL 133 133 MANGANESE (BY SIMILARITY).
 CC FT METAL 137 137 MANGANESE (BY SIMILARITY).
 CC FT NON_TER 138 138
 CC SQ SEQUENCE 138 AA; 14976 MW; BBB5A3C41B1F028 CRC64;
 CC -----
 CC Query Match 1.5%; Score 7; DB 1; Length 138;
 CC Best Local Similarity 100.0%; Pred. No. 26;
 CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC QY 209 SLGQKLL 215
 CC DB 105 SLGQKLL 111
 CC | | | | | | |
 CC
 CC RESULT 10
 CC Y335_TREPA STANDARD; PRT; 180 AA.
 CC AC O83355;
 CC DT 20-AUG-2001 (Rel. 40, Created)
 CC DT 20-AUG-2001 (Rel. 40, Last sequence update)
 CC DT 20-AUG-2001 (Rel. 40, Last annotation update)
 CC DE HYPOTHETICAL PROTEIN TP0335.
 CC GN TP0335.
 CC OS Treponema pallidum.
 CC OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
 CC OX NCBI_TaxID=160;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC RC STRAIN=NICHOLS;
 CC RX MEDLINE=98332770; PubMed=9665876;
 CC RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
 CC Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
 CC Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
 CC Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
 CC McDonald L., Artisach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
 CC Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
 CC Venter J.C.;
 CC "Complete genome sequence of Treponema pallidum, the syphilis
 CC Spirochete";
 CC RL Science 281:375-388(1998).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF001213; AAC65328.1; -.
 CC TIGR; TP0335; -.
 CC KW Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 37 59 POTENTIAL.
 FT TRANSMEM 128 147 POTENTIAL.
 SQ SEQUENCE 180 AA; 19218 MW; 5BB9C80BB334B48A CRC64;
 CC -----
 CC Query Match 1.5%; Score 7; DB 1; Length 180;
 CC Best Local Similarity 100.0%; Pred. No. 33;
 CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC QY 191 GEVLSVL 197
 CC DB 113 GEVLSVL 119
 CC | | | | | | |
 CC
 CC RESULT 11
 CC GCAL_HUMAN STANDARD; PRT; 200 AA.
 CC ID GCAL_HUMAN
 CC AC P43080; O9NU14;
 CC DT 01-NOV-1995 (Rel. 32, Created)
 CC DT 01-NOV-1995 (Rel. 32, Last sequence update)
 CC DT 20-AUG-2001 (Rel. 40, Last annotation update)
 CC DE GUANYLYL CYCLASE ACTIVATING PROTEIN 1 (GCAP 1) (GUANYLATE CYCLASE
 CC ACTIVATOR 1A).
 CC GN GUCAL1 OR GUCAL OR GCAP1 OR GCAP.
 CC OS Homo sapiens (Human).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CC OX NCBI_TaxID=9606;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC RC TISSUE=Retina;
 CC RX MEDLINE=95074147; PubMed=7983048;
 CC RA Subbaraya I., Ruiz C.C., Helekar B.S., Zhao X., Gortzyca W.A.,
 CC Pettenati M.J., Rao P.N., Palczewski K., Baehr W.;
 CC "Molecular characterization of human and mouse photoreceptor
 CC guanylate cyclase-activating protein (GCAP) and chromosomal
 CC localization of the human gene";
 CC RL J. Biol. Chem. 269:31080-31089(1994).
 CC [2]
 CC SEQUENCE FROM N.A.
 CC RA Sehra H.;
 CC RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC [3]
 CC RP VARIANT COD3 CYS-98.
 CC RX MEDLINE=98087579; PubMed=9425234;
 CC RA Payne A.M., Downes S.M., Bessant D.A.R., Taylor R., Holder G.E.,
 CC Warren M.J., Bird A.C., Bhattacharya S.S.;
 CC "A mutation in guanylate cyclase activator 1A (GUCAL1) in an autosomal
 CC dominant cone dystrophy pedigree mapping to a new locus on chromosome
 CC 6p21.1";
 CC RL Hum. Mol. Genet. 7:273-277(1998).
 CC [4]
 CC CHARACTERIZATION OF VARIANTS COD3 LEU-49 AND CYS-98.
 CC RX MEDLINE=20562747; PubMed=11108966;
 CC RA Sokal I., Li N., Verlinde C.L., Haeseleer F., Baehr W., Palczewski K.,
 CC "Ca(2+)-binding proteins in the retina: from discovery to etiology of
 CC human disease";
 CC RL Biochim. Biophys. Acta 1498:233-251(2000).
 CC [5]
 CC VARIANTS COD3 LEU-49 AND CYS-98.
 CC RX MEDLINE=20581605; PubMed=11146732;
 CC RA Downes S.M., Holder G.E., Fitzke F.W., Payne A.M., Warren M.J.,
 CC Bhattacharya S.S., Bird A.C.;
 CC "Autosomal dominant cone and cone-rod dystrophy with mutations in the
 CC guanylate cyclase activator 1A gene-encoding guanylate cyclase
 CC activating protein-1";
 CC RL Arch. Ophthalmol. 119:96-105(2001).
 CC -!- FUNCTION: STIMULATES GUANYLYL CYCLASE 1 (GC1) WHEN FREE CALCIUM
 CC IONS CONCENTRATION IS LOW AND INHIBITS GC1 WHEN FREE CALCIUM IONS
 CC CONCENTRATION IS ELEVATED. THIS CA(2+)-SENSITIVE REGULATION OF GC
 CC IS A KEY EVENT IN RECOVERY OF THE DARK STATE OF ROD PHOTORECEPTORS
 CC FOLLOWING LIGHT EXPOSURE.
 CC -!- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED.


```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89359361; PubMed=2527850;
RA Ra C., Jouvin M.H.E., Kinet J.-P.;
RT "Complete structure of the mouse mast cell receptor for IgE (Fc
RT epsilon RI) and surface expression of chimeric receptors (rat-mouse-
human) on transfected cells.";
RL J. Biol. Chem. 264:15323-15327(1989).
CC -!- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULINS EPSILON. HIGH
CC AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
CC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL
CC ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
CC RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
CC ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
CC -!- SUBUNIT: Tetramer of an alpha chain, a beta chain, and two
CC disulfide linked gamma chains.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J05019; AAA37601.1; -
CC PIR; B34342; B34342.
CC MGD; MGI:95495; M5441.
KW IgE-binding protein; Receptor; Transmembrane.
FT DOMAIN 1 51 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 52 71 POTENTIAL.
FT DOMAIN 72 89 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 90 109 POTENTIAL.
FT DOMAIN 110 122 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 123 142 POTENTIAL.
FT DOMAIN 143 171 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 172 191 POTENTIAL.
FT DOMAIN 192 235 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 235 AA; 25963 MW; 1C2D6BF195738168 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 IAAGTGI 105
Db 131 IAAGTGI 137
|||||

RESULT 14
ATP6_MYCLE
ID ATP6_MYCLE STANDARD; PRT; 251 AA.
AC P45829;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ATP SYNTHASE A CHAIN (EC 3.6.1.34) (PROTEIN 6).
GN ATPB OR ML1139.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith D.R., Robison K.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RL

```

```

RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jacels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.-A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 405:1007-1011(2001).
CC -!- FUNCTION: KEY COMPONENT OF THE PROTON CHANNEL; IT MAY PLAY A
CC SUBUNIT ROLE IN THE TRANSLLOCATION OF PROTONS ACROSS THE MEMBRANE.
CC -!- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC HAS THREE MAIN SUBUNITS: A, B AND C.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE ATPASE A CHAIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U15186; AAA63110.1; -
CC EMBL; AL583920; CAC31520.1; -
CC Leproma; ML1139; -
DR InterPro; IPR000568; ATP_synt_A.
DR Pfam; PF00119; ATP_synt_A; 1.
DR PRINTS; PR00123; ATPASEA; 1.
DR PROSITE; PS00449; ATPASEA; 1.
KW Hydrogen ion transport; CF(0); Transmembrane; Complete proteome.
FT TRANSMEM 28 48 POTENTIAL.
FT TRANSMEM 84 104 POTENTIAL.
FT TRANSMEM 130 150 POTENTIAL.
FT TRANSMEM 154 174 POTENTIAL.
FT TRANSMEM 192 212 POTENTIAL.
FT TRANSMEM 220 240 POTENTIAL.
SQ SEQUENCE 251 AA; 27558 MW; BCCF2AE7FC158DF9 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 251;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 203 TTEVLSS 209
Db 119 TTEVLSS 125
|||||

RESULT 15
PSA7_CAPEL
ID PSA7_CAPEL STANDARD; PRT; 253 AA.
AC Q95005;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROBABLE PROTEASOME SUBUNIT ALPHA TYPE 7 (EC 3.4.99.46).
GN C36B1.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RL Lennard N.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.

```

CC -!- FUNCTION: THE PROTEASOME IS A MULTICATALYTIC PROTEINASE COMPLEX
CC WHICH IS CHARACTERIZED BY ITS ABILITY TO CLEAVE PEPTIDES WITH ARG,
CC PHE, TYR, LEU, AND GLU ADJACENT TO THE LEAVING GROUP AT NEUTRAL OR
CC SLIGHTLY BASIC PH. THE PROTEASOME HAS AN ATP-DEPENDENT PROTEOLYTIC
CC ACTIVITY.
CC -!- PATHWAY: INVOLVED IN AN ATP/UBIQUITIN-DEPENDENT NON-LYSOSOMAL
CC PROTEOLYTIC PATHWAY.
CC -!- SUBUNIT: THE PROTEASOME IS COMPOSED OF AT LEAST 15 NON IDENTICAL
CC SUBUNITS WHICH FORM A HIGHLY ORDERED RING-SHAPED STRUCTURE (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY T1A; ALSO KNOWN AS THE
CC PROTEASOME A-TYPE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z80215; CAB02269.1; -
DR HSSP; P25156; LPMA.
DR MEROPS; T01.974; -
DR WormPep; C36B1.4; CE05371.
DR InterPro; IPR001353; Proteasome.
DR Pfam; PF00227; proteasome_A.
DR PROSITE; PS00388; PROTEASOME_A; 1.
KW Proteasome; Hydrolase; Protease.
SQ SEQUENCE 253 AA; 28239 MW; 345078EDBE908EDC CRC64;

Query Match 1.5%; Score 7; DB 1; Length 253;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 LAFAGLS 221
Db 72 LAFAGLS 78
|||||

RESULT 16
CTRL_HUMAN STANDARD; PRT; 264 AA.
ID CTRL_HUMAN
AC P40313;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CHYMOTRYPSIN-LIKE PROTEASE CTRL-1 PRECURSOR (EC 3.4.21.-).
GN CTRL OR CTRL1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94093544; PubMed=8268911;
RA Larsen F., Solheim J., Kristensen T., Kolsto A.B., Prydz H.;
RT "A tight cluster of five unrelated human genes on chromosome
RT 16q22.1.";
RL Hum. Mol. Genet. 2:1589-1595(1993).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

CC EMBL; X71874; CAA50710.1; -
DR EMBL; X71877; CAA50711.1; -
DR HSSP; P00763; IDPO.
DR MEROPS; S01.256; -
DR MIM; 118888; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Trypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT PROPEP 19 33 ACTIVATION PEPTIDE (POTENTIAL).
FT CHAIN 34 264 CHYMOTRYPSIN-LIKE PROTEASE CTRL-1.
FT ACT_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 121 121 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 214 214 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CAROHYD 114 114 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 19 141 BY SIMILARITY.
FT DISULFID 60 76 BY SIMILARITY.
FT DISULFID 155 220 BY SIMILARITY.
FT DISULFID 187 201 BY SIMILARITY.
FT DISULFID 210 239 BY SIMILARITY.
SQ SEQUENCE 264 AA; 28002 MW; 3F629F02FA6DDDFB4 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 264;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 LSLVLLG 226
Db 8 LSLVLLG 14
|||||

RESULT 17
CY1_RHOCA STANDARD; PRT; 279 AA.
ID CY1_RHOCA
AC P08501; P07058;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE CYTOCHROME C1 PRECURSOR.
GN PETC.
OS Rhodobacter capsulatus (Rhodospseudomonas capsulata).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Rhodobacter.
OX NCBI_TaxID=1061;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88011223; PubMed=2821268;
RA Davidson E., Balda F.;
RT "Primary structure of the bcl complex of Rhodospseudomonas capsulata.
RT Nucleotide sequence of the pet operon encoding the Rieske cytochrome
RT b, and cytochrome c1 apoproteins.";
RL J. Mol. Biol. 195:13-24(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=GA;
RX MEDLINE=86136096; PubMed=3004982;
RA Gabellini N., Sebald W.;
RT "Nucleotide sequence and transcription of the fbc operon from
RT Rhodospseudomonas sphaeroides. Evaluation of the deduced amino acid
RT sequences of the Fes protein, cytochrome b and cytochrome c1.";
RL Eur. J. Biochem. 154:569-579(1986).
RN [3]
RP CORRECTION OF ORGANISM GIVEN IN REF.2.
RX MEDLINE=88011233; PubMed=2821272;


```
RL Unpublished observations (JAN-1995).
CC -1- FUNCTION: MAY PLAY A ROLE IN THE REPAIR OF ENDOGENOUS ALKYLATION
CC DAMAGE.
CC -1- SIMILARITY: BELONGS TO THE RADA FAMILY.
CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A FRAMESHIFT HAD TO
CC BE INTRODUCED IN POSITION 109 TO PRODUCE THIS ORF.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M85181; -; NOT_ANNOTATED_CDS.
DR PIR; A37753; A37753.
DR STYGene; SGI0444; rada.
KW DNA repair; ATP-binding; DNA-binding.
FT NON_TER 1
SQ SEQUENCE 298 AA; 32119 MW; E85668E5B9D8AA04 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 298;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 LSVFDDL 164
Db 292 LSVFDDL 298
|||||

RESULT 20
ICIA_AERSA STANDARD; PRT; 299 AA.
AC P07773;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE CHROMOSOME INITIATION INHIBITOR (ORIC REPLICATION INHIBITOR).
GN ICIA.
OS Aeromonas salmonicida.
OC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;
OC Aeromonas.
OX NCBI_TaxID=645;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCIMB 1102;
RA Swift S., Kariyshev A.V., Fish L., Durant E.L., Winson M.K.,
RA Williams P., Macintyre S., Stewart G.S.A.B.;
RA Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: SPECIFIC INHIBITOR OF CHROMOSOMAL INITIATION OF
CC REPLICATION IN VITRO. BINDS THE THREE 13-MERS IN THE ORIGIN
CC (ORIC) TO BLOCK INITIATION OF REPLICATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U65741; AAB70016.1; -
DR InterPro; IPR000847; HTH_Lysr.
DR Pfam; PF00126; HTH_1; 1.
DR PRINTS; PR00039; HTHLYSR.
DR PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
KW DNA-binding; Transcription regulation.
FT DNA_BIND 21 40 H-T-H MOTIF (POTENTIAL).
SQ SEQUENCE 299 AA; 33410 MW; 5EEBECA44CC5A68D CRC64;
```

```
Query Match 1.5%; Score 7; DB 1; Length 299;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 210 LQOKLLA 216
Db 62 LQOKLLA 68
|||||

RESULT 21
YEDA_ECOLI STANDARD; PRT; 306 AA.
ID YEDA_ECOLI AC P09185;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN YEDA.
GN YEDA OR B1959 OR Z3050 OR ECS2697.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=89345179; PubMed=2527357;
RA Hanck T., Gerwin N., Fritz H.-J.;
RT "Nucleotide sequence of the dcu locus of Escherichia coli K12."
RL Nucleic Acids Res. 17:5844-5844(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251358; PubMed=9097040;
RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,
RA Saito N., Sampei G., Seki Y., Sivasubraman S., Tagami H.,
RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 40.1-50.0 min region on the linkage map."
RL DNA Res. 3:379-392(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Postbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoukis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
```

```

RA Lida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
CC -!- SIMILARITY: BELONGS TO THE EMA TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X13330; CAA31708.1; -
CC EMBL; AE000287; AAC75025.1; -
CC EMBL; D90835; BAA15786.1; -
CC EMBL; AE003418; AAG56973.1; -
CC EMBL; AF002559; BAB36120.1; -
CC PIR; JS0266; JS0266.
CC EcoGene; EGI1141; yedA.
CC InterPro; IPR000620; DUF6.
CC Pfam; PF00892; DUF6; 2.
KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
FT TRANSMEM 7 27 POTENTIAL.
FT TRANSMEM 37 57 POTENTIAL.
FT TRANSMEM 68 88 POTENTIAL.
FT TRANSMEM 94 114 POTENTIAL.
FT TRANSMEM 126 146 POTENTIAL.
FT TRANSMEM 149 169 POTENTIAL.
FT TRANSMEM 174 194 POTENTIAL.
FT TRANSMEM 207 227 POTENTIAL.
FT TRANSMEM 240 260 POTENTIAL.
FT TRANSMEM 270 290 POTENTIAL.
FT SEQUENCE 306 AA; 32194 MW; 6F0D63E8090ACA64 CRC64;
Query Match 1.5%; Score 7; DB 1; Length 306;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 223 VLLGTGL 229
DB 254 VLLGTGL 260
|||||||
RESULT 22
ID EFTS_BARQU STANDARD; PRT; 307 AA.
AC Q9XCM5;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE ELONGATION FACTOR TS (EF-TS).
GN TSF.
OS Bartonella quintana (Rochalimaea quintana).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bartonellaceae; Bartonella.
OX NCBI_TaxID=803;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OK90-268;
RA Marston E.L., Eldieief S., Regnery R.L.;
RT "Cloning and characterization of elongation factor-ts (EF-ts) gene
RT from Bartonella quintana.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ASSOCIATES WITH THE EF-TU.GDP COMPLEX AND INDUCES THE
CC EXCHANGE OF GDP TO GTP, IT REMAINS BOUND TO THE AMINOACYL-TRNA.
CC EF-TU.GTP COMPLEX UP TO THE GTP HYDROLYSIS STAGE ON THE RIBOSOME.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE EF-TS FAMILY.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF138286; AAD39149.1; -
CC InterPro; IPR001816; EF_TS.
CC InterPro; IPR003037; TS-N.
CC Pfam; PF00889; EF_TS; 1.
CC Pfam; PF02094; TS-N; 1.
CC PROSITE; PS01126; EF_TS_1; FALSE_NEG.
CC PROSITE; PS01127; EF_TS_2; 1.
CC Elongation factor; Protein biosynthesis.
FT SITE 79 82 INVOLVED IN MG++ ION DISLOCATION FROM EF-
FT TU (BY SIMILARITY).
SQ SEQUENCE 307 AA; 32453 MW; 6C78271FB457FC6B CRC64;
Query Match 1.5%; Score 7; DB 1; Length 307;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 404 VSASFYP 410
DB 109 VSASFYP 115
|||||||
RESULT 23
ID CBSB_SULSO STANDARD; PRT; 314 AA.
AC P58030;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE CYTOCHROME B558/566 SUBUNIT B.
GN CBSB OR SSO2802 OR C48.010.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweyez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Senses C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL512979; CAC23868.1; -
CC EMBL; AE006874; AAK42915.1; -
KW Electron transport; Transmembrane; Complete proteome.
FT TRANSMEM 47 67 POTENTIAL.
FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 102 122 POTENTIAL.
FT TRANSMEM 127 147 POTENTIAL.

```

FT TRANSMEM 155 175 POTENTIAL.
 FT TRANSMEM 186 206 POTENTIAL.
 FT TRANSMEM 233 253 POTENTIAL.
 FT TRANSMEM 264 284 POTENTIAL.
 SQ SEQUENCE 314 AA; 35030 MW; 7D14E8F01922ECA4 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 314;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 ISVGLLS 48
 |||||
 Db 189 ISVGLLS 195

RESULT 24
 YZ34_AQUAE STANDARD; PRT; 318 AA.
 AC O66423;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHETICAL PROTEIN AA34.
 GN AA34
 OS Aquifex aeolicus.
 OG Plasmid ecel.
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus.";
 RL Nature 392:353-358(1998).
 CC -!- SIMILARITY: STRONG, TO A. AEOLICUS AA07 AND AA11.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE000667; AAC07975.1; -
 KW Hypothetical protein; Plasmid; Complete proteome.
 SQ SEQUENCE 318 AA; 37360 MW; EDB125C719DE49A3 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 318;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 TEVLSSL 210
 |||||
 Db 17 TEVLSSL 23

RESULT 25
 FABH_MYCTU STANDARD; PRT; 335 AA.
 AC O06399;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE 3-OXOACYL-[ACYL-CARRIER-PROTEIN] SYNTHASE III (EC 2.3.1.41) (BETA-
 DE KETOACYL-ACP SYNTHASE III) (KAS III).
 GN FABH OR RV0533C OR MT0557 OR MTCY25D10.12C.

OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekai F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CATALYZES THE CONDENSATION REACTION OF FATTY ACID
 CC SYNTHESIS BY THE ADDITION TO AN ACYL ACCEPTOR OF TWO CARBONS
 CC FROM MALONYL-ACP. KAS III CATALYZES THE FIRST CONDENSATION
 CC REACTION WHICH INITIATES FATTY ACID SYNTHESIS AND MAY THEREFORE
 CC PLAY A ROLE IN GOVERNING THE TOTAL RATE OF FATTY ACID PRODUCTION.
 CC POSSESSES BOTH ACETOACETYL-ACP SYNTHASE AND ACETYL TRANSACYLASE
 CC ACTIVITIES (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ACYL-[ACYL-CARRIER PROTEIN] + MALONYL-[ACYL-
 CC CARRIER PROTEIN] = 3-OXOACYL-[ACYL-CARRIER PROTEIN] + CO(2) +
 CC [ACYL-CARRIER PROTEIN]
 CC -!- PATHWAY: FATTY ACID BIOSYNTHESIS.
 CC -!- SIMILARITY: BELONGS TO THE FABH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; 295558; CAB08984.1; -
 DR EMBL; AE006954; AAK44780.1; -
 DR TIGR; MT0557; -
 DR Tuberculist; RV0533c; -
 KW Fatty acid biosynthesis; Transferase; Acyltransferase;
 KW Multifunctional enzyme; Complete proteome.
 FT ACT_SITE 122 122 BY SIMILARITY.
 FT ACT_SITE 258 258 BY SIMILARITY.
 FT ACT_SITE 289 289 BY SIMILARITY.
 SQ SEQUENCE 335 AA; 34872 MW; 6573BE1FAE5BCFB6 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 335;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 SVGLLSV 49
 |||||
 Db 12 SVGLLSV 18

RESULT 26

```

CMST_CRIGR
ID CMST_CRIGR STANDARD; PRT; 336 AA.
AC O08520;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CMP-SIALIC ACID TRANSPORTER (CMP-SIA-TR) (CMP-SA-TR).
GN SLC35A1.
OS Cricetulus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97454302; PubMed=9310377;
RA Eckhardt M., Gerardy-Schahn R.;
RT "Molecular cloning of the hamster CMP-sialic acid transporter.";
RL Eur. J. Biochem. 248:187-192(1997).
CC -!- FUNCTION: TRANSPORT CMP-SIALIC ACID FROM THE CYTOSOL INTO GOLGI
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. GOLGI.
CC -!- SIMILARITY: BELONGS TO THE NUCLEOTIDE-SUGAR TRANSPORTER FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; 271268; CAA95855.1; .
DR MGD; MGI:1345622; SLC35a1.
KW Transport; Sugar transport; Transmembrane; Golgi stack.
FT TRANSMEM 31 POTENTIAL.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 140 160 POTENTIAL.
FT TRANSMEM 175 195 POTENTIAL.
FT TRANSMEM 213 233 POTENTIAL.
FT TRANSMEM 238 258 POTENTIAL.
FT TRANSMEM 288 308 POTENTIAL.
SQ SEQUENCE 336 AA; 36608 MW; 209D8C52A67D8F9 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 336;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 LISVGLL 47
Db 57 LISVGLL 63

RESULT 27
CMST_MOUSE
ID CMST_MOUSE STANDARD; PRT; 336 AA.
AC Q61420;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CMP-SIALIC ACID TRANSPORTER (CMP-SIA-TR) (CMP-SA-TR).
GN SLC35A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96353858; PubMed=8755516;
RA Eckhardt M., Muehlenhoff M., Bethé A., Gerardy-Schahn R.;
RT "Expression cloning of the Golgi CMP-sialic acid transporter.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:7572-7576(1996).
CC -!- FUNCTION: TRANSPORT CMP-SIALIC ACID FROM THE CYTOSOL INTO GOLGI

EMBL; Y12074; CAA72794.1; .
KW Transport; Sugar transport; Transmembrane; Golgi stack.
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 140 160 POTENTIAL.
FT TRANSMEM 175 195 POTENTIAL.
FT TRANSMEM 213 233 POTENTIAL.
FT TRANSMEM 238 258 POTENTIAL.
FT TRANSMEM 288 308 POTENTIAL.
SQ SEQUENCE 336 AA; 36608 MW; 209D8C52A67D8F9 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 336;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 LISVGLL 47
Db 57 LISVGLL 63

RESULT 28
YHNCM_ECOLI
ID YHNCM_ECOLI STANDARD; PRT; 375 AA.
AC P46442;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN YHCM.
GN YHCM OR B3232 OR Z4591 OR ECS4105.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Escherichia.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MGI1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [3]

```


CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -!- SIMILARITY: STRONG, TO M.TUBERCULOSIS RV3630.
CC -!- CAUTION: REF.2 SEQUENCES DIFFER FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 173.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: 279701; CAB02024.1; -
CC EMBL: AE007023; AAK45825.1; ALT_FRAME.
CC EMBL: AE007023; AAK45826.1; ALT_FRAME.
CC TIGR: MT1558; -
CC TIGR: MT1560; -
CC TubercuList: Rv1510; -
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 35 55 POTENTIAL.
FT TRANSMEM 60 80 POTENTIAL.
FT TRANSMEM 112 132 POTENTIAL.
FT TRANSMEM 144 164 POTENTIAL.
FT TRANSMEM 185 205 POTENTIAL.
FT TRANSMEM 209 229 POTENTIAL.
FT TRANSMEM 242 262 POTENTIAL.
FT TRANSMEM 274 294 POTENTIAL.
FT TRANSMEM 313 333 POTENTIAL.
FT TRANSMEM 359 379 POTENTIAL.
FT TRANSMEM 384 404 POTENTIAL.
FT TRANSMEM 408 428 POTENTIAL.
SQ SEQUENCE 432 AA; 44293 MW; 96ECD0C48814D7F6 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 432;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 SVGLLSV 49
Db 143 SVGLLSV 149

Search completed: January 31, 2002, 08:18:43
Job time: 439 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2002, 08:10:24 ; Search time 57.51 Seconds
(without alignments)
1195.409 Million cell updates/sec

Title: US-09-713-098-2
Perfect score: 470
Sequence: 1 MGIWTSGTDFISLWEIYVS.....LPVLKMITRKKOMDNASADKS 470

Scoring table: OLIGO

Searched: 473505 seqs, 146272329 residues

Word size : 0

Total number of hits satisfying chosen parameters: 473505

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Listing first 50 summaries

```
Database :
SPTREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhnc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query		DB	ID	Description
	Score	Match Length			
1	470	100.0	470	4	Q9h295 homo sapien
2	18	3.8	367	11	Q9d619 mus musculu
3	8	1.7	96	8	Q9g2x0 myxine glut
4	8	1.7	101	8	Q63922 myxine glut
5	8	1.7	176	8	Q63517 echinostoma
6	8	1.7	177	2	Q9K312 vibrio chol
7	8	1.7	338	2	Q9x724 streptomyce
8	8	1.7	461	10	Q9C968 arabidopsis
9	8	1.7	506	10	Q9aw0 cryza sativ
10	8	1.7	736	5	Q9wly0 drosophila
11	8	1.7	1025	3	Q9uux9 neurospora
12	8	1.7	1285	2	Q9wxu3 thermotoga
13	7	1.5	65	5	Q9bpd5 conus ventr
14	7	1.5	78	2	Q9klm0 neisseria m
15	7	1.5	97	8	Q03355 lutzomyia w
16	7	1.5	105	2	Q9znc8 salmonella
17	7	1.5	106	2	Q46528 bacteroides
18	7	1.5	110	2	Q52835 bordetella
19	7	1.5	111	10	Q9lwl1 cryza sativ

20	7	1.5	117	2	Q9A6U4	Q9A6U4	caulobacter
21	7	1.5	120	12	Q02124	Q02124	poplar moss
22	7	1.5	122	10	Q9FVX3	Q9FVX3	arabidopsis
23	7	1.5	126	2	P72643	P72643	mycobacteri
24	7	1.5	129	2	P94916	P94916	mycobacteri
25	7	1.5	131	4	O60531	O60531	homo sapien
26	7	1.5	142	2	Q9KWX3	Q9KWX3	streptococc
27	7	1.5	150	1	O27777	O27777	methanobact
28	7	1.5	153	5	Q37790	Q17790	caenorhabdi
29	7	1.5	155	2	Q36966	Q33696	streptococc
30	7	1.5	158	2	Q9KD19	Q9kd19	bacillus ha
31	7	1.5	160	2	O31659	O51659	borrelia bu
32	7	1.5	162	2	Q9JUD7	Q9juu7	neisseria m
33	7	1.5	164	10	Q9LFK3	Q9lfm3	arabidopsis
34	7	1.5	168	3	Q62336	Q06236	saccharomyc
35	7	1.5	176	8	O63510	Q63510	echinostoma
36	7	1.5	176	8	O63513	O63513	echinostoma
37	7	1.5	176	8	O63514	O63514	echinostoma
38	7	1.5	182	5	Q18388	Q18388	caenorhabdi
39	7	1.5	184	10	Q9SSX3	Q9ssx3	barbula unc
40	7	1.5	188	2	Q9RU36	Q9ru36	deinococcus
41	7	1.5	190	5	Q18519	Q18519	caenorhabdi
42	7	1.5	220	2	Q9REN9	Q9ren9	zymomonas m
43	7	1.5	220	2	Q9FD46	Q9fd46	streptococc
44	7	1.5	220	5	Q9W4T0	Q9w4t0	drosohila
45	7	1.5	227	2	Q82891	O82891	escherichia
46	7	1.5	250	5	O02212	O02212	caenorhabdi
47	7	1.5	252	4	Q9R3K4	Q9y3k4	homo sapien
48	7	1.5	259	2	Q36754	Q56754	bergeyella
49	7	1.5	262	2	Q9FD44	Q9fd44	streptococc
50	7	1.5	262	5	O9XK67	Q9xx67	caenorhabdi

ALIGNMENTS

```

RESULT      1
Q9H295
Q9H295      PRELIMINARY;      PRT;      470 AA.
ID
Q9H295;
AC
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE DC-SPECIFIC TRANSMEMBRANE PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP
RA SEQUENCE FROM N.A.
RA Hartgers F.C., Visser J.L.M., Looman M.W.G., van Zoelen C.,
RT Huffin C., Figdor C.G., Adema G.-J.;
RA "DC-STAMP, a novel multimembrane-spanning molecule preferentially
RT expressed by dendritic cells.";
RL Eur. J. Immunol. 0:0-0(2001).
DR EMBL; AF305068; ARG39167.1; -.
DR InterPro; IPR001211; PIP_A2.
DR PROSITE; PS00118; PA2_HIS; UNKNOWN_1.
KW Transmembrane.
SQ SEQUENCE      470 AA; 53392 MW; EA28858FD2C7560C CRC64;

```

Query Match	100.0%	Score 470;	DB 4;	Length 470;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 470;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

1	MG1WTSGTDIFLSWIIYIYSPSPGWMDFTHQILGVCCCLVALISVGLLSVAACWFLPSIIA	60	QY
1	MG1WTSGTDIFLSWIIYIYSPSPGWMDFTHQILGVCCCLVALISVGLLSVAACWFLPSIIA	60	Db
61	AAASWIIITCVLLCCSHKRCFILLVFLSCGLREGRNALIAAGTGVILGHVENIFINFKG	120	QY
61	AAASWIIITCVLLCCSHKRCFILLVFLSCGLREGRNALIAAGTGVILGHVENIFINFKG	120	Db

QY 121 LLDGMCNLRKSFSEHFPLKKYIEAIONWYGLATPLSVFDDLVSNQTLAVSLFSPSH 180
|||||
Db 121 LLDGMCNLRKSFSEHFPLKKYIEAIONWYGLATPLSVFDDLVSNQTLAVSLFSPSH 180
QY 181 VLEAQLNDSGEVLSVLYQWATTEVSSLGOKLAFAGLSVLGTLGFLMKRFLGPCGW 240
|||||
Db 181 VLEAQLNDSGEVLSVLYQWATTEVSSLGOKLAFAGLSVLGTLGFLMKRFLGPCGW 240
QY 241 KYENIYITROFVQFDERERHQPCVLPVLPNKEERKRYVIPTFWPTPKERNLGLFFLPI 300
|||||
Db 241 KYENIYITROFVQFDERERHQPCVLPVLPNKEERKRYVIPTFWPTPKERNLGLFFLPI 300
QY 301 LIHLCIWLFAAVDYLLYRILFVSQFQSLPGFVEVHLKLGHEKQGTQDIHDSFNSIV 360
|||||
Db 301 LIHLCIWLFAAVDYLLYRILFVSQFQSLPGFVEVHLKLGHEKQGTQDIHDSFNSIV 360
QY 361 FEPNCIPKPFLLSETWVPLSVILLILVLMGLSSILMQLKILVSASFYPSVERKRIQYL 420
|||||
Db 361 FEPNCIPKPFLLSETWVPLSVILLILVLMGLSSILMQLKILVSASFYPSVERKRIQYL 420
QY 421 HAKLLKRSKQPLGEVKRRRLSLYTKTHFWLPVLKMKIRKKOMDASADKS 470
|||||
Db 421 HAKLLKRSKQPLGEVKRRRLSLYTKTHFWLPVLKMKIRKKOMDASADKS 470
RESULT 2
ID Q9D619 PRELIMINARY; PRT; 367 AA.
AC Q9D619;
DT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
DE 4833414I07RIK PROTEIN.
GN 4833414I07RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=HEAD;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weltz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasedawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK014697; BAB29508.1; -;
DR MGD; MGI:1923016; 4833414I07RIK.
SQ SEQUENCE 367 AA; 42104 MW; 964ADCF855B543B1 CRC64;

Query Match 3.8%; Score 18; DB 11; Length 367;
Best Local Similarity 100.0%; Pred. No. 5.2e-09;
Matches, 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 87 LSCGLREGNALTAAAGTG 104

Db 87 LSCGLREGNALTAAAGTG 104
|||||
RESULT 3
Q9G2X0 PRELIMINARY; PRT; 96 AA.
ID Q9G2X0;
AC Q9G2X0;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE NADH DEHYDROGENASE SUBUNIT 4L.
GN NADH4L.
OS Myxine glutinosa (Atlantic hagfish).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniiformes;
OC Myxiniidae; Myxiniinae; Myxine.
OX NCBI_TaxID=7769;
RN [1]
RP SEQUENCE FROM N.A.
RA Delarbre C., Rasmussen A.S., Arnason U., Gachelin G.;
RL "Complete sequence of the mitochondrial DNA of Myxine glutinosa.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ404477; CAC20657.1; -;
DR InterPro; IPR003214; Mit_NADHub_oxidredctse_4L.
DR InterPro; IPR001133; Oxidored_q2.
DR Pfam; PF00420; Oxidored_q2; 1.
DR ProDom; PD000359; Mit_NADHub_oxidredctse_4L; 1.
KW Mitochondrion.
SQ SEQUENCE 96 AA; 10501 MW; 862D807596D6CB26 CRC64;
Query Match 1.7%; Score 8; DB 8; Length 96;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 218 AGLSLVLL 225
|||||
Db 69 AGLSLVLL 76
RESULT 4
O63922 PRELIMINARY; PRT; 101 AA.
ID O63922;
AC O63922;
DT 01-AUG-1998 (TremBLrel. 07, Created)
DT 01-AUG-1998 (TremBLrel. 07, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE NADH DEHYDROGENASE SUBUNIT 4L.
GN NADH4L.
OS Myxine glutinosa (Atlantic hagfish).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniiformes;
OC Myxiniidae; Myxiniinae; Myxine.
OX NCBI_TaxID=7769;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98210228; PubMed=9541532;
RA Rasmussen A.S.G., Janke A., Arnason U.;
RT "The mitochondrial DNA molecule of the hagfish (Myxine glutinosa) and
RT vertebrate phylogeny.";
RL J. Mol. Evol. 46:382-388(1998).
DR EMBL; Y15191; CAA75490.1; -;
DR InterPro; IPR001133; Oxidored_q2.
DR InterPro; IPR003214; Mit_NADHub_oxidredctse_4L.
DR Pfam; PF00420; Oxidored_q2; 1.
DR ProDom; PD000359; Mit_NADHub_oxidredctse_4L; 1.
KW Mitochondrion.
SQ SEQUENCE 101 AA; 11000 MW; 712C16F448C97E56 CRC64;
Query Match 1.7%; Score 8; DB 8; Length 101;
Best Local Similarity 100.0%; Pred. No. 11;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 AGLSLVLL 225
| | | | | | | |

Db 74 AGLSLVLL 81

RESULT 5
063517 PRELIMINARY; PRT; 176 AA.

AC 063517;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE NICOTINAMIDE ADENINE DINUCLEOTIDE DEHYDROGENASE SUBUNIT 1 (FRAGMENT).
GN NDI.
OS Echinostoma sp. Rat-Ad.
OG Mitochondrion.
OC Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Neodermata;
OC Trematoda; Digenea; Echinostomida; Echinostomata; Echinostomatidae;
OC Echinostoma.
OX NCBI_TaxID=68345;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RAT-AD (AUSTRALIA);
RX MEDLINE=98220121; PubMed=9559367;
RA Morgan J.A., Blair D.;
RT "Mitochondrial NDI gene sequences used to identify echinostome
RT isolates from Australia and New Zealand.";
RL Int. J. Parasitol. 28:493-502(1998).
DR EMBL; AF026290; AAC15966.1; -;
DR InterPro; IPR001694; Resp_chain_NADH_DH1.
DR Pfam; PF00146; NADHdh; 2.
KW Mitochondrion.
FT NON_TER 1 176
FT NON_TER 176 176
SQ SEQUENCE 176 AA; 19279 MW; C91374B6B8747172 CRC64;

Query Match 1.7%; Score 8; DB 8; Length 176;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 LLVFLSCG 90
| | | | | | | |

Db 45 LLVFLSCG 52

RESULT 6
Q9KSI2 PRELIMINARY; PRT; 177 AA.

AC Q9KSI2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE HYPOTHETICAL PROTEIN VC1274.
GN VC1274.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.K., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";

RL Nature 406:477-483(2000).
DR EMBL; AE004206; AAF94433.1; -;
DR TIGR; VC1274; -;
KW Complete proteome.
SQ SEQUENCE 177 AA; 20019 MW; EEB34A82C5F525C8 CRC64;

Query Match 1.7%; Score 8; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 390 LGLLSSIL 397
| | | | | | | |

Db 18 LGLLSSIL 25

RESULT 7
Q9X7Z4 PRELIMINARY; PRT; 338 AA.

AC Q9X7Z4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PUTATIVE SECRETED PROTEIN.
GN SC6G10.08C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K., Harris D.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RA Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL049497; CAB39862.1; -;
DR InterPro; IPR000064; NLPC_P60.
DR Pfam; PF00877; NLPC_P60; 1.
SQ SEQUENCE 338 AA; 35703 MW; 3BCF02D3151449C8 CRC64;

Query Match 1.7%; Score 8; DB 2; Length 338;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 VGLLSVAA 51
| | | | | | | |

Db 20 VGLLSVAA 27

RESULT 8
Q9C968 PRELIMINARY; PRT; 461 AA.

AC Q9C968;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PUTATIVE CYCLIN.

GN F516.12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLOMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway T.H., Dewar K.,
RA Dunn P., Egu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana.";
RL Nature 408:816-820(2000).
CC -1- STIMILARITY: TO CYCLIN FAMILY.
DR EMBL; AC018848; AAG52439.1;
DR InterPro; IPR000553; Cyclin.
DR Pfam; PF00134; Cyclin; 1.
DR SMART; SM00385; CYCLIN; 2.
KW Cell cycle; Cell division; Cyclin.
SQ SEQUENCE 461 AA; 51636 MW; 7282354BBF7BA241 CRC64;

Query Match 1.7%; Score 8; DB 10; Length 461;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 FLPSIIIA 61
| | | | |
Db 373 FLPSIIIA 380

RESULT 9
Q9AWW0 PRELIMINARY; PRT; 506 AA.
AC Q9AWW0
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
GN PUTATIVE GLYCOPOLIPID ALPHA-MANNOSYLTRANSFERASE.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sakai T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:PO013G02.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002908; BAB32949.1;
KW Transferase; Glycosyltransferase.
SQ SEQUENCE 506 AA; 57706 MW; C8D438CCBD2A462F CRC64;

Query Match 1.7%; Score 8; DB 10; Length 506;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 380 LSVILLIL 387
| | | | |
Db 377 LSVILLIL 384
| | | | |
RESULT 10
Q9WLY0 PRELIMINARY; PRT; 736 AA.
AC Q9WLY0
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE CG3499 PROTEIN.
GN CG3499
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aebavani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busan D.A., Butler H.C., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003459; AAF46922.1;
DR FlyBase; FBgn0034792; CG3499.
DR InterPro; IPR003593; AAA.
DR InterPro; IPR003960; AAA_sub.
DR InterPro; IPR003959; AAA_subfam.
DR InterPro; IPR000642; Peptidase_M41.
DR Pfam; PF00004; AAA; 1.
DR Pfam; PF01434; Peptidase_M41; 1.

DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS00674; AAA; 1.
 SQ SEQUENCE 736 AA; 80650 MW; C976701830B8BC52 CRC64;

Query Match 1.7%; Score 8; DB 5; Length 736;
 Best Local Similarity 100.0%; Pred. No. 61;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 440 LSLYLTKI 447
 |||||
 Db 477 LSLYLTKI 484

RESULT 11
 Q9UUX9 PRELIMINARY; PRT; 1025 AA.
 AC Q9UUX9;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE PUTATIVE CALCIN P-TYPE ATPASE (FRAGMENT).
 GN PMR1.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Benito B., Garciladeblas B., Rodriguez-Navarro A.;
 RT "Calcium and sodium ATPases in Neurospora crassa."
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ243518; CAB65296.1; -
 DR HSSP; P04191; LEUL.
 DR InterPro; IPR001757; E1-E2.ATPase.
 DR InterPro; IPR001454; Hydrolase.
 DR InterPro; IPR000661; Na_H_K_ATPase.
 DR Pfam; PF00122; E1-E2_ATPase; 1.
 DR Pfam; PF00702; Hydrolase; 1.
 DR Pfam; PF00690; Na_K_ATPase_N; 1.
 DR PRINTS; PR00119; CATATPASE.
 DR PRINTS; PR00121; NAKATPASE.
 DR PROSITE; PS00154; ATPASE_E1_E2; UNKNOWN_1.
 FT NON_TER 1025 1025
 SQ SEQUENCE 1025 AA; 110822 MW; 1F7C7681D0D44964 CRC64;

Query Match 1.7%; Score 8; DB 3; Length 1025;
 Best Local Similarity 100.0%; Pred. No. 82;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 AGLSLVLL 225
 |||||
 Db 820 AGLSLVLL 827

RESULT 12
 Q9WXU3 PRELIMINARY; PRT; 1285 AA.
 AC Q9WXU3;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE COME PROTEIN, PUTATIVE.
 GN TM0088.
 OS Thermotoga maritima.
 OC Bacteria; Thermotogales; Thermotoga.
 OX NCBI_TaxID=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,

RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RL genome sequence of *Thermotoga maritima*."
 RL Nature 398:323-329(1999).
 DR EMBL; AE001695; AAD35182.1; -
 DR TIGR; TM0088; -
 DR InterPro; IPR000016; Bac_GSPproteins.
 DR InterPro; IPR001993; Mitoch_carrier.
 DR Pfam; PF00263; Bac_GSPproteins; 1.
 DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 1285 AA; 145209 MW; 057435F821FB0EA5 CRC64;

Query Match 1.7%; Score 8; DB 2; Length 1285;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 383 ILLILVML 390
 |||||
 Db 7 ILLILVML 14

RESULT 13
 Q9BPD5 PRELIMINARY; PRT; 65 AA.
 AC Q9BPD5;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE CONOTOXIN PRECURSOR.
 OS Conus ventricosus.
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=117992;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21105969; PubMed=11158371;
 RA Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,
 RA Fainzilber M.;
 RT "Mechanisms for Evolving Hypervariability: The Case of Conopeptides."
 RL Mol. Biol. Evol. 18:120-131(2001).
 DR EMBL; AF214999; AAC60427.1; -
 SQ SEQUENCE 65 AA; 6933 MW; 42E526E3753A22FA CRC64;

Query Match 1.5%; Score 7; DB 5; Length 65;
 Best Local Similarity 100.0%; Pred. No. 70;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 391 GLLSSIL 397
 |||||
 Db 50 GLLSSIL 56

RESULT 14
 Q9K1M0 PRELIMINARY; PRT; 78 AA.
 AC Q9K1M0;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE HYPOTHETICAL PROTEIN NMB0091.
 GN NMB0091.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 RN [1]
 RP SEQUENCE FROM N.A.

```

RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; Pubmed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citrone H., Clark E.B.,
RA Cotton M.D., Uterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizsa M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58.";
RL Science 287:1809-1815(2000).
DR EMBL; AE002368; AAF40553.1; -.
DR TIGR; NMB0091; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 78 AA; 8311 MW; 62EB649BBEDC4F1C CRC64;

Query Match 1.5%; Score 7; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 322 FSVSKQF 328
Db 51 FSVSKQF 57
|||||||

RESULT 15
O03355 PRELIMINARY; PRT; 97 AA.
AC O03355;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE CYTOCHROME B (FRAGMENT).
GN CYTB.
OS Lutzomyia whitmani.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Psychodoidea;
OC Psychodidae; Lutzomyia.
OX NCBI_TaxID=55416;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NORTH EAST.
RA Ready P.D., Day J.C., de Souza A.A., Rangel E.F., Davies C.R.;
RA Bull. Entomol. Res. 87:187-195(1997).
CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B/B6 FAMILY.
DR EMBL; U80966; BAB58257.1; -.
DR InterPro; IPR00179; Cyt_b56.
DR Pfam; PF00032; cytochrome_b_c1; 1.
KW Electron transport; Heme; Mitochondrion; Respiratory chain;
FT Transmembrane.
FT NON_TER 1
SQ SEQUENCE 97 AA; 11250 MW; C38BFF6A312ED497 CRC64;

Query Match 1.5%; Score 7; DB 8; Length 97;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 295 LFFLPIL 301
Db 20 LFFLPIL 26
|||||||

```

```

RESULT 16
O9ZHC8 PRELIMINARY; PRT; 105 AA.
AC O9ZHC8;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DE 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
DE HYPOTHETICAL 11.1 KDA PROTEIN.
GN ORF105.
OS Salmonella typhimurium.
OG Plasmid pMG101.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RA Gupta A., Matsui K., Lo J.-F., Silver S.;
RT "Molecular basis for resistance to silver cations in Salmonella.";
RL Nat. Med. 0:0-0(1999).
DR EMBL; AF067954; AAD11752.1; -.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 105 AA; 11118 MW; F311E57E12D70EC2 CRC64;

Query Match 1.5%; Score 7; DB 2; Length 105;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 IYGLATP 157
Db 63 IYGLATP 69
|||||||

RESULT 17
Q46528 PRELIMINARY; PRT; 106 AA.
AC Q46528;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE HYPOTHETICAL PROTEIN (FRAGMENT).
OS Bacteroides nodosus (Dichelobacter nodosus).
OC Bacteria; Proteobacteria; gamma subdivision; Cardiobacteriaceae;
OC Dichelobacter.
OX NCBI_TaxID=870;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A198 (VCS 1001);
RX MEDLINE=97090407; Pubmed=8936315;
RA La Fontaine S., Rood J.I.;
RT "Organization of ribosomal RNA genes from the footrot pathogen
RT Dichelobacter nodosus.";
RL Microbiology 142:889-899(1996).
DR EMBL; U26680; AAC13872.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 106 AA; 12179 MW; 5A5E2C94B178B05D CRC64;

Query Match 1.5%; Score 7; DB 2; Length 106;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 210 LGQKLLA 216
Db 54 LGQKLLA 60
|||||||

RESULT 18
O52835 PRELIMINARY; PRT; 110 AA.
ID O52835

```

AC 052835;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE BCR PROTEIN (FRAGMENT).
 GN BCR.
 OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
 OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
 OC Bordetella.
 OX NCBI_TaxID=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BB1015;
 RX MEDLINE=98132383; PubMed=9473041;
 RA Pradel E., Guiso N., Loch C.;
 RT "Identification of Alcr, an AraC-type regulator of alcaligin
 RT siderophore synthesis in Bordetella bronchiseptica and Bordetella
 RT pertussis".
 RL J. Bacteriol. 180:871-880(1998).
 DR EMBL; AJ000061; CAA03894.1; -;
 FT NON_TER 110 110
 SQ SEQUENCE 110 AA; 11661 MW; F13A8E7A1C6ED6F5 CRC64;

Query Match 1.5%; Score 7; DB 2; Length 110;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 388 VMLGLLS 394
 |||||
 Db 15 VMLGLLS 21

RESULT 19

Q9LW11
 ID Q9LW11 PRELIMINARY; PRT; 111 AA.
 AC Q9LW11;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE EST AU062706 (C30225) CORRESPONDS TO A REGION OF THE PREDICTED GENE.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
 RT clone: P0675A05".
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF002071; BAA95878.1; -;
 SQ SEQUENCE 111 AA; 12557 MW; 30502903E6ADC2CF CRC64;

Query Match 1.5%; Score 7; DB 10; Length 111;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 VALISVG 45
 |||||
 Db 10 VALISVG 16

RESULT 20

Q9A6U4
 ID Q9A6U4 PRELIMINARY; PRT; 117 AA.
 AC Q9A6U4;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

DE HYPOTHETICAL PROTEIN CCI1989.
 GN CCI1989.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 OX NCBI_TaxID=69394;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21173698; PubMed=11259647;
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA DeBoy R.T., Dodson R.J., Burkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus".
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL; AE005872; AAK23964.1; -;
 DR TIGR; CCI1989; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 117 AA; 12189 MW; 820990801EB59E84 CRC64;

Query Match 1.5%; Score 7; DB 2; Length 117;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 223 VLLGTGL 229
 |||||
 Db 62 VLLGTGL 68

RESULT 21

Q02124
 ID Q02124 PRELIMINARY; PRT; 120 AA.
 AC Q02124;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE HYPOTHETICAL PROTEIN IN HELICASE GENE 3' REGION (FRAGMENT).
 OS Poplar mosaic virus (isolate ATCC PV275) (PMV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Carlavirus.
 OX NCBI_TaxID=31709;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92333281; PubMed=1629709;
 RA Henderson J., Gibbs M.J., Edwards M.L., Clarke V.A., Gardner K.A.,
 RA Cooper J.I.;
 RT "Partial nucleotide sequence of poplar mosaic virus RNA confirms its
 RT classification as a carlavirus".
 RL J. Gen. Virol. 73:1887-1890(1992).
 DR EMBL; D13364; BAA02626.1; -;
 DR InterPro; IPR001896; Plant_vir_prot.
 DR Pfam; PF01307; Plant_vir_prot; 1.
 DR ProDom; PD001561; Plant_vir_prot; 1.
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 120 AA; 12543 MW; ED85E0E736E117FD CRC64;

Query Match 1.5%; Score 7; DB 12; Length 120;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 AGLSLVL 224
 |||||
 Db 30 AGLSLVL 36

RESULT 22

Q9FVX3
 ID Q9FVX3 PRELIMINARY; PRT; 122 AA.

```

AC Q9FVX3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE HYPOTHETICAL 13.0 KDA PROTEIN.
GN F2P24.6.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Renning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome 1 BAC F2P24 genomic sequence.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC078898; AG29199.1; -
KW Hypothetical protein.
SQ SEQUENCE 122 AA; 13049 MW; 2F823CC8AE2CE721 CRC64;

Query Match 1.5%; Score 7; DB 10; Length 122;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 ALIAAGT 103
DB 78 ALIAAGT 84

RESULT 23
P72643 PRELIMINARY; PRT; 126 AA.
AC P72643;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE THIOREDOXIN M.
GN TRX OR SLL1057.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugiyura M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsumoto A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; D90899; BAA16645.1; -
DR HSP; P00274; ITP7.
DR InterPro; IPR000063; Thioired.
DR Pfam; PF00085; Thioired; 1.
DR PRINTS; PR00421; THIOREDOXIN.
KW Complete proteome.
SQ SEQUENCE 126 AA; 14633 MW; 87C92DE42EDFF5D CRC64;

Query Match 1.5%; Score 7; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 AVSLFSP 178
DB 120 AVSLFSP 126

```

```

RESULT 24
P94916 PRELIMINARY; PRT; 129 AA.
AC P94916;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SUPEROXIDE DISMUTASE (FRAGMENT).
GN SOD.
OS Mycobacterium abscessus
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=36809;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19977T;
RA Domenech P., Jimenez M.S., Menendez C., Bull T.J., Samper S.,
RA Manrique A., Garcia M.J.;
RL Int. J. Syst. Bacteriol. 0:0-0(0).
CC -!- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).
CC -!- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
DR EMBL; U86082; AAB46719.1; -
DR HSP; P17670; IDS.
DR InterPro; IPR001189; SOD_MI.
DR Pfam; PF00081; sodfe; 1.
DR ProDom; PD000475; SOD_MI; 1.
KW Oxidoreductase.
FT NON_TER 1
FT NON_TER 129
SQ SEQUENCE 129 AA; 13976 MW; 2BF96B3B035677ED CRC64;

Query Match 1.5%; Score 7; DB 2; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 SLGQKLL 215
DB 105 SLGQKLL 111

RESULT 25
O60531 PRELIMINARY; PRT; 131 AA.
AC O60531;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DE 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE ANTIGEN NY-CO-16.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLORECTAL CARCINOMA;
RA Scanlan M.J., Chen Y.T., Williamson B., Gure A.O., Stockert E.,
RA Gordon J.D., Toreci O., Sahin V., Pfreundschuh M., Old L.J.;
RL Int. J. Cancer 0:0-0(1998).
DR EMBL; AF039694; AAC18043.1; -
SQ SEQUENCE 131 AA; 14639 MW; 6FD7A76394169581 CRC64;

Query Match 1.5%; Score 7; DB 4; Length 131;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 LPLNKEE 273

```

```

Db 118 LPLNKEE 124
|||||
RESULT 26
O9KWY3 PRELIMINARY; PRT; 142 AA.
AC O9KWY3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE RIBONUCLEOTIDE REDUCTASE-LIKE (NRD-LIKE) PROTEIN.
GN NRD.
OS Streptococcus equisimilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=119602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=25287;
RX MEDLINE=20123454; PubMed=10660058;
RA Geyer A., Schmidt K.H.;
RT "Genetic organisation of the M protein region in human isolates of
RT group C and G streptococci: two types of multigene regulator-Like
RT (mgrC) regions.";
RL Mol. Gen. Genet. 262:965-974 (2000).
DR EMBL; Y18363; CAB70606.1; -.
DR InterPro; IPR003707; NrdI.
DR Pfam; PF02555; NrdI; 1.
SQ SEQUENCE 142 AA; 15991 MW; B2F50660B307C470 CRC64;

Query Match 1.5%; Score 7; DB 2; Length 142;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 439 RLSLYLT 445
|||||
Db 21 RLSLYLT 27

RESULT 27
O27777 PRELIMINARY; PRT; 150 AA.
AC O27777;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE PROTEIN DISULPHIDE ISOMERASE.
GN MTH1745.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanothermobacter.
OX NCBI_TaxID=145262;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DELTA H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., DuBois J.,
RA Aldredge T., Bashirzaden R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jivani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;
RA "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155 (1997).
DR EMBL; AE000930; AAB86215.1; -.
DR InterPro; IPR000063; Thiored.
DR Pfam; PF00085; thiored; 1.
DR PROSITE; PS00194; THIOREDOXIN; UNKNOWN_1.
KW Isomerase; Complete proteome.

```

```

SQ SEQUENCE 150 AA; 16939 MW; E0672E64B7E340EB CRC64;

Query Match 1.5%; Score 7; DB 1; Length 150;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 AGLSLVL 224
|||||
Db 19 AGLSLVL 25

RESULT 28
Q17790 PRELIMINARY; PRT; 153 AA.
AC Q17790;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE C07E3.9 PROTEIN.
GN C07E3.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Matthews P.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38 (1994).
DR EMBL; Z49908; CAA90098.1; -.
DR HSSP; P00593; 4BP2
DR InterPro; IPR001211; PLP_A2.
DR Pfam; PF00088; phoslip; 1.
DR ProDom; PD000303; PLP_A2; 1.
DR SMART; SM00085; PA2c; 1.
DR PROSITE; PS00119; PA2_ASP; UNKNOWN_1.
DR PROSITE; PS00118; PA2_HIS; UNKNOWN_1.
SQ SEQUENCE 153 AA; 17490 MW; EBDDB0E059CEA329 CRC64;

Query Match 1.5%; Score 7; DB 5; Length 153;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 FILLVFL 87
|||||
Db 3 FILLVFL 9

RESULT 29
O33696 PRELIMINARY; PRT; 155 AA.
AC O33696;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)
DE 15.5 KDA PROTEIN PRECURSOR.

```

```
GN PGS.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TYPE M25/174;
RA Gubbe K., Schmidt K.H.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; X92371; CAA63115.1; -.
KW Signal.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 155 POTENTIAL.
SQ SEQUENCE 155 AA; 17250 MW; B3CDDFFEBE8F46CA CRC64;

Query Match 1.5%; Score 7; DB 2; Length 155;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 221 SLVLLGT 227
Db 21 SLVLLGT 27

RESULT 30
Q9KD19
ID Q9KD19 PRELIMINARY; PRT; 158 AA.
AC Q9KD19;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE BH1400 PROTEIN.
GN BH1400.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AF001511; BAB05119.1; -.
KW Complete proteome.
SQ SEQUENCE 158 AA; 17032 MW; E17BA1F9630F39B5 CRC64;

Query Match 1.5%; Score 7; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 295 LFFLPIL 301
Db 110 LFFLPIL 116
```

Search completed: January 31, 2002, 08:18:16
Job time: 472 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2002, 08:07:44 ; Search time 56.83 Seconds
(without alignments)
612.607 Million cell updates/sec

Title: US-09-713-098-2
Perfect score: 470
Sequence: 1 MGWTSGTDFLSLWEIYVS.....LPVLKMKRKKQMDMASADKS 470

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 522463 seqs, 74073290 residues

Word size : 0
Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : A_Geneseq_1101.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	470	100.0	470	22	AAE02638 Human dendritic ce
2	283	60.2	352	22	AAE02638 Human gene 16 enco
3	250	53.2	257	22	AAE03317 Human gene 20 enco
4	250	53.2	257	22	AAE03317 Human gene 16 enco
5	250	53.2	257	22	AAE03317 Human gene 16 enco
6	250	53.2	257	22	AAE03317 Human gene 16 enco
7	239	50.9	292	21	AAE03317 Human secreted pro
8	176	37.4	291	22	AAE03317 Human gene 16 enco
9	7	1.5	35	20	AAE03317 Human 5' EST secre
10	7	1.5	45	15	AAE03317 Characteristic pro
11	7	1.5	54	20	AAE03317 Barley germin N-te

12	7	1.5	57	21	AAE03317 Mouse Interleukin-
13	7	1.5	75	21	AAE03317 Human signal pepti
14	7	1.5	114	21	AAE03317 Arabidopsis thalia
15	7	1.5	114	21	AAE03317 Arabidopsis thalia
16	7	1.5	121	21	AAE03317 Arabidopsis thalia
17	7	1.5	122	21	AAE03317 Arabidopsis thalia
18	7	1.5	127	20	AAE03317 Protein involved i
19	7	1.5	150	22	AAE03317 Novel bone marrow
20	7	1.5	152	21	AAE03317 Arabidopsis thalia
21	7	1.5	157	22	AAE03317 Human protein sequ
22	7	1.5	162	21	AAE03317 Neisseria meningit
23	7	1.5	163	21	AAE03317 Arabidopsis thalia
24	7	1.5	164	21	AAE03317 Arabidopsis thalia
25	7	1.5	187	21	AAE03317 Arabidopsis thalia
26	7	1.5	192	21	AAE03317 Human pancreatic c
27	7	1.5	198	20	AAE03317 Histamine binding
28	7	1.5	200	20	AAE03317 Protein which is s
29	7	1.5	217	22	AAE03317 Human gene 7 enco
30	7	1.5	222	22	AAE03317 C glutamicum prote
31	7	1.5	224	22	AAE03317 Corynebacterium gl
32	7	1.5	236	20	AAE03317 Colon cancer assoc
33	7	1.5	264	21	AAE03317 Human serine prote
34	7	1.5	264	21	AAE03317 Mouse serine prote
35	7	1.5	278	17	AAE03317 ILTV short repeat
36	7	1.5	280	22	AAE03317 Human secreted pro
37	7	1.5	285	22	AAE03317 Human secreted pro
38	7	1.5	293	21	AAE03317 Human secreted pro
39	7	1.5	309	20	AAE03317 Human secreted pro
40	7	1.5	349	22	AAE03317 Human gene 7 enco
41	7	1.5	350	20	AAE03317 Homo sapiens secre
42	7	1.5	352	21	AAE03317 Arabidopsis thalia
43	7	1.5	355	18	AAE03317 Mycobacterium tube
44	7	1.5	355	18	AAE03317 Mycobacterium tube
45	7	1.5	355	19	AAE03317 M. tuberculosis im
46	7	1.5	355	19	AAE03317 M. tuberculosis im
47	7	1.5	355	19	AAE03317 Mycobacterium tube
48	7	1.5	355	20	AAE03317 M. tuberculosis an
49	7	1.5	355	20	AAE03317 M. tuberculosis re
50	7	1.5	355	20	AAE03317 Mycobacterium spec

ALIGNMENTS

RESULT 1
AAE02638
ID AAE02638 standard; Protein; 470 AA.
XX
AC AAE02638;
XX
DT 06-AUG-2001 (first entry)
XX
DE Human dendritic cell specific transmembrane protein (DC-STAMP).
XX
KW Human; dendritic cell specific transmembrane protein; DC-STAMP;
KW forensic science; therapy; abnormal physiology; allergic condition;
KW asthma; cancer; autoimmune disease; diabetes mellitus;
KW drug screening.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 35..51
FT Domain /label= Transmembrane_domain_1
FT Domain 57..75
FT Domain /label= Transmembrane_domain_2
FT Domain 96..114
FT Domain /label= Transmembrane_domain_3
FT Domain 144..162
FT Domain /label= Transmembrane_domain_4
FT Modified-site 168..170
FT /note= "Asn is N-glycosylated"
FT Modified-site 187..189

allergies, neurological disorders (e.g., Alzheimer's disease, Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders, angiogenic disorders, kidney disorders, gastrointestinal disorders, pregnancy-related disorders, endocrine disorders, and infections. The proteins can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues, to identify their cognate ligands or binding partners, and in chemotaxis, and can be used as a food additive or preservative to modify storage properties. Antibodies specific for a protein of the invention can be used in alleviating symptoms associated with the disorders mentioned above, and in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The present sequence represents a human secreted protein of the invention.

AA	Sequence	352 AA;
SQ		

Query Match 60.2%; Score 283; DB 22; Length 352;
Best Local Similarity 100.0%; Pred. No. 3.6e-264;

Qy	27	MDFTQHLGVCCILVALISVGLLSVAAACWFLPSITIAAASWIITCVLLCCSKHARCFILLVF	86
Db	1	mdfqiqlgvceclvalisvglsvaaacwflpsiaaaaswiitcvllccskharcfillvf	60
Qy	87	LSCSLREGRNALIAAGTGVILVGHVENIHFNFKGLDGMTCNLRKSFSTHPPLKKYIE	146
Db	61	lscslregrnaliaaagtgvilghveni hf n f k l d g m t c n l r k s f s t h p p l k k y i e	120
Qy	147	AIQWYGLATPLSVFDDLVSWNQTLAVSLFSPSHVLVLAQLNSKSGEVLVLYQMATTVE	206
Db	121	a i q w y g l a t p l s v f d d l v s w n q t l a v s l f s p s h v l v l a q l n s k s g e v l v l y q m a t t v e	180
Qy	207	LSSLGOKLAFAGLSVLVLTGTLFPMKRFIIGPCGWKYNENIYTRQFVQFDERRHQORPCV	266
Db	181	l s s l g k l a f a g l s v l v l t g t l f m k r f i g p c g w k y n e n i y t r q f v q f d e r e r h q o r p c v	240
Qy	267	LPUNKERRKYIIPFTWPTPKRNKLGFLFFLPILIHLCIWWL	309
Db	241	lpunkerrkyvliptfwptpkernknglffipilhlclwvl	283

RESULTS

AAE03917

ID AAE03917 standard; protein: 257 AA.

AA
AC AAF03917:

AA	
DT	09-AUG-2001 (first entry)

Human gene 20 encoded secreted protein HMDJ14, SEO ID NO:80.

Human; secreted protein; proliferative disorder; cancer; tumour;
fetal abnormality; developmental abnormality; haematopoietic disorder;
immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
inflammation; allergy; neurological disorder; Alzheimer's disease;
Parkinson's disease; cognitive disorder; schizophrenia; asthma;
skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
cardiovascular disorder; angiogenic disorder; kidney disorder;
gastrointestinal disorder; pregnancy-related disorder;
endocrine disorder; infection; wound healing; vulvuary;
cell culture; chemotaxis; food additive; gene therapy;
binding partner identification.

AA
OS
Homo sapiens.

XX	Key	Location/Qualifiers
FH		

Key	Location
FT	1..26
Peptide	

```
FT
FT peptide
1..20
/label= signal_peptide
```

FI		/label
FT	Protein	27..257

FT	/note="Mature secreted protein"
----	---------------------------------

PN WO200077022-A1.

21-DEC-2000.

01-JUN-2000: 2000WO-US15136.

PR 11-JUN-1999; 99US-0138629.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Komatsoulis GA;

DR WPI; 2001-367020/38.

DR N-PSDB; AAD08364.

PT Nucleic acids encoding 50 human secreted polypeptides, useful for preventing, diagnosing and/or treating diseases, e.g. Parkinson's disease, botulism, cancers and Scimitar syndrome -

PS Claim 11: Page 535-536: 614pp: English:

AA080345-AA080394 represent cDNAs corresponding to 50 human secreted protein genes and AA03989-AA03947 represent the proteins they encode. AA03948-AA03996 represent human secreted protein fragments or variants. The genes and their secreted proteins are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Pathological conditions can be diagnosed by determining the amount of the new protein in a sample or by determining the presence of mutations in the new genes. Specific uses are described for each of the 50 genes, based on the tissues in which they are most highly expressed, and include developing products for the diagnosis or treatment of proliferative disorders, cancer, tumours, foetal and developmental abnormalities, haematopoietic disorders, diseases of the immune system, AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation, allergies, neurological disorders (e.g., Alzheimer's disease, Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders, angio-genic disorders, kidney disorders, gastrointestinal disorders, pregnancy-related disorders, endocrine disorders, and infections. The proteins can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues, to identify their cognate ligands or binding partners, and in chemotaxis, and can be used as a food additive or preservative to modify storage properties. Antibodies specific for a protein of the invention can be used in alleviating symptoms associated with the disorders mentioned above, and in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The present sequence represents a human secreted protein of the invention.

XX	Sequence	257 AA:
SO		

Query Match 53.2%: Score 250: DB 22: Length 257:

Query Match 55.2%, Score 250, DB 22, Length 251,
Best Local Similarity 100.0%, Pred. No. 1.7e-232;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OV 27 MDFIOHLGVCCCLVALISVGLLSVAACWFELPSIIAAAAASWIIITCVLLCCSKHARCFILLVF 86

1 mdfiqhlayccclvalisvqllsvaacwflpsiaaaaaawiiitcvllccskharcfilllyf 60

0v 87 I.SCGI.REGRNALJAAGTGIVILGHVENIFHNFKGLLDGMTCNLRAKSFSIHFFLLKKYIE 146

nb 67 | sca | rearna | aagtat | i | qhven | f | hnfka | | damt | cn | raks | fs | ib | fol | k | kvie | 120

147 A TQWYIGIATPI SVFDDI VSWNOTI AVSLFSPSHVIFAOLNDSKGEVI SVIYOMATTEV 206

nb 121 a1cw1vq1atn1cyfdd1vswqat1avs1fnsbhy1eag1ndskqov1sv1ymat1t1ey 180

207 ISSIGOKILAEAGLSIVLGTGLFMKREI.GPCCWKYENTYITROFVOEDERERHOORPCV 266

|||||
181 lsslqgkllafaglsllvllgtlglfmkrlgpcgkwyeniytrqfvqderhqrpcv 240
QY 267 LPLNKEERRK 276
|||||
Db 241 lplnkeerrk 250

RESULT 4
AAB87399
ID AAB87399 standard; Protein: 257 AA.
XX AC AAB87399;
XX DF 22-MAY-2001 (first entry)
XX DE Human gene 16 encoded secreted protein HMADJ14, SEQ ID NO:140.
XX DE Human; secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angiogenic disorder; kidney disorder;
KW gastrointestinal disorder; infection; pregnancy-related disorder;
KW endocrine disorder; wound healing; vulnery;
KW cell culture; chemotaxis; food additive;
KW binding partner identification.
XX OS Homo sapiens.
XX PN WO200118022-A1.
XX PD 15-MAR-2001.
XX PF 31-AUG-2000; 2000WO-US24008.
XX PR 03-SEP-1999; 99US-0152315.
XX PR 03-SEP-1999; 99US-0152317.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Ni J, Baker KP, Birse CE, Fiscella M, Komatsoulis GA, Rosen CA;
PI Soppet DR, Young PE, Ebner R, Duan DR, Olsen HS, Lafleur DW;
PI Moore PA, Shi Y, Wei Y, Florence KA;
XX WPI: 2001-203081/20.
XX N-PSDB; AAF91915.

Nucleic acid molecules encoding human secreted proteins, used in
preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
Parkinson's diseases and cancers -

Claim 11: Page 567-568; 607pp; English.

AAAF91858-AAF91929 represent cDNAs corresponding to 52 human secreted
protein genes, and AAB87342-AAB87413 represent the proteins they encode.
AAB87414-AAB87454 represent human secreted protein fragments. The genes
and their corresponding secreted proteins are useful for preventing,
treating or ameliorating medical conditions, e.g., by protein or gene
therapy. Pathological conditions can be diagnosed by determining the
amount of the new protein in a sample or by determining the presence of
mutations in the new genes. Specific uses are described for each of the
52 genes, based on the tissues in which they are most highly expressed,
and include developing products for the diagnosis or treatment of
proliferative disorders, cancer, tumours, foetal and developmental
abnormalities, haematopoietic disorders, diseases of the immune system,
AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
allergies, neurological disorders (e.g., Alzheimer's disease, asthma,
Parkinson's disease), cognitive disorders, schizophrenia, atherosclerosis,
skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
cardiovascular disorders, angiogenic disorders, kidney disorders,

CC gastrointestinal disorders, pregnancy-related disorders, endocrine
CC disorders, and infections. The proteins can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin aging due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues, to identify their
CC cognate ligands or binding partners, and in chemotaxis, and can be used
CC as a food additive or preservative to modify storage properties.
CC Antibodies specific for a protein of the invention can be used in
CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunosorbent assay (ELISA). The present sequence represents a human
XX secreted protein of the invention.
SQ Sequence 257 AA;

Query Match 53.2%; Score 250; DB 22; Length 257;
Best Local Similarity 100.0%; Pred. No. 1.7e-232; Indels 0; Gaps 0;
Matches 250; Conservative 0; Mismatches 0;

QY 27 MDFIQHLGVCCLVALISVGLLSVAACWFLPSIIIAAASWIIITCVLLCCSKHARCFILLVF 86
|||||
Db 1 mdfiqhlgvccivalisvlgllsvaacwflpsiaaaswiitcvllccskharcfillvf 60

QY 87 LSCGLREGRNALIACTGIVILGHVENIFHNFKGLDGMTCNIRAKSFSTHFPLKKYIE 146
|||||
Db 61 lscglregrnallaagtgvilghvenifhnfkglldgmcniraksfshfplkkyyie 120

QY 147 ATQWLYGLATPLSVFDDLYSWNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 206
|||||
Db 121 atqwyglatpplsvfddlyswnqtlavslfshvleaqindskgevlsvlyqmattev 180

QY 207 LSSLGOKLLAFAGLSLVLLGTGLFMKRFGLPCGCKWYENIYTRQFVQDERRHOORPCV 266
|||||
Db 181 lsslqgkllafaglsllvllgtlglfmkrlgpcgkwyeniytrqfvqderhqrpcv 240

QY 267 LPLNKEERRK 276
|||||
Db 241 lplnkeerrk 250

RESULT 5
AAB87400
ID AAB87400 standard; Protein: 257 AA.
XX AC AAB87400;
XX DF 22-MAY-2001 (first entry)
XX DE Human gene 16 encoded secreted protein HMADJ14, SEQ ID NO:141.
XX DE Human; secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angiogenic disorder; kidney disorder;
KW gastrointestinal disorder; infection; pregnancy-related disorder;
KW endocrine disorder; wound healing; vulnery;
KW cell culture; chemotaxis; food additive;
KW binding partner identification.
XX OS Homo sapiens.
XX PN WO200118022-A1.
XX PD 15-MAR-2001.
XX PF 31-AUG-2000; 2000WO-US24008.
XX PR 03-SEP-1999; 99US-0152315.
XX PR 03-SEP-1999; 99US-0152317.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Ni J, Baker KP, Birse CE, Fiscella M, Komatsoulis GA, Rosen CA;

PI Soppet DR, Young PE, Ebner R, Duan DR, Olsen HS, Lafleur DW;

PI Moore PA, Shi Y, Wei Y, Florence KA;

XX WPI: 2001-203081/20.

DR N-PSDB; AAF91916.

XX Nucleic acid molecules encoding human secreted proteins, used in

PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and

PT Parkinson's diseases and cancers -

XX Claim 11; Page 568-569; 607pp; English.

XX AAF91858-AAF91929 represent cDNAs corresponding to 52 human secreted

CC protein genes, and AAB87342-AAB87413 represent the proteins they encode.

CC AAB87414-AAB87454 represent human secreted protein fragments. The genes

CC and their corresponding secreted proteins are useful for preventing,

CC treating or ameliorating medical conditions, e.g., by protein or gene

CC therapy. Pathological conditions can be diagnosed by determining the

CC amount of the new protein in a sample or by determining the presence of

CC mutations in the new genes. Specific uses are described for each of the

CC 52 genes, based on the tissues in which they are most highly expressed,

CC and include developing products for the diagnosis or treatment of

CC proliferative disorders, cancer, tumours, foetal and developmental

CC abnormalities, haematopoietic disorders, diseases of the immune system,

CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,

CC allergies, neurological disorders (e.g., Alzheimer's disease,

CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,

CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,

CC cardiovascular disorders, angiogenic disorders, kidney disorders,

CC gastrointestinal disorders, pregnancy-related disorders, endocrine

CC disorders, and infections. The proteins can also be used to aid wound

CC healing and epithelial cell proliferation, to prevent skin aging due to

CC sunburn, to maintain organs before transplantation, for supporting cell

CC culture of primary tissues, to regenerate tissues, to identify their

CC cognate ligands or binding partners, and in chemotaxis, and can be used

CC as a food additive or preservative to modify storage properties.

CC Antibodies specific for a protein of the invention can be used in

CC alleviating symptoms associated with the disorders mentioned above, and

CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked

CC immunosorbent assay (ELISA). The present sequence represents a human

CC secreted protein of the invention.

XX Sequence 257 AA;

XX

Query Match 53.2%; Score 250; DB 22; Length 257;

Best Local Similarity 100.0%; Pred. NO. 1.7e-232;

Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 MDTFHLGVCCLLVALLSVGLLSVAACWFLPSIITAAASWIIITCVLLCCSHKRCFILLVF 86

Db 1 mdfihlgvccllvallsvglslsvaacwflpsiaaaswiitcvllccshkrcflllvf 60

QY 87 LSCGLREGRNALIAAGTGVILGHVENIFHNFKGLDGMTCNLRKSFSTHFPLKKYIE 146

Db 61 lscglregrnaliaagtgvilghvenifhnfkgldgmtcnlraksfshfplkkylie 120

QY 147 AIQWYGLATPLSFVDDLVSWNQT LAVSLFSPSHVLAEQLNDSKGEVLSVLYQMATTTEV 206

Db 121 aiqwyglatplsvfddlvswnqtlavslfspshvlleaqlndskgevlsvlyqmatteev 180

QY 207 LSSLGQKLLAFAGLSVLVLTGTFEMRFLGPCWKNYIYITQVFQFDERHRHQPCV 266

Db 181 lsslgqkllafaglsvlvltgtfemrflgpcwknyniyitqvfqfderhrhqgrpcv 240

QY 267 LPLNKERRK 276

Db 241 lplnkeerkr 250

RESULT 6

AAB87454

ID AAB87454 standard; Protein; 257 AA.

XX AC AAB87454;

XX 22-MAY-2001 (first entry)

XX Human gene 16 encoded secreted protein fragment.

XX Human; secreted protein; proliferative disorder; cancer; tumour;

KW foetal abnormality; developmental abnormality; haematopoietic disorder;

KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;

KW inflammation; allergy; neurological disorder; Alzheimer's disease;

KW Parkinson's disease; psoriasis; sepsis; diabetes; atherosclerosis;

KW skin disorder; cognitive disorder; schizophrenia; asthma;

KW cardiovascular disorder; angiogenic disorder; kidney disorder;

KW gastrointestinal disorder; pregnancy-related disorder;

KW endocrine disorder; infection; wound healing; vunerary;

KW cell culture; chemotaxis; food additive;

KW binding partner identification.

OS Homo sapiens.

XX WO200118022-A1.

XX 15-MAR-2001.

XX 31-AUG-2000; 2000WO-US24008.

XX 03-SEP-1999; 99US-0152315.

XX 03-SEP-1999; 99US-0152317.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ni J, Baker KP, Birse CE, Fiscella M, Komatsoulis GA, Rosen CA;

PI Soppet DR, Young PE, Ebner R, Duan DR, Olsen HS, Lafleur DW;

PI Moore PA, Shi Y, Wei Y, Florence KA;

XX WPI: 2001-203081/20.

XX Nucleic acid molecules encoding human secreted proteins, used in

PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and

PT Parkinson's diseases and cancers -

XX Disclosure; Page 55; 607pp; English.

XX AAF91858-AAF91929 represent cDNAs corresponding to 52 human secreted

CC protein genes, and AAB87342-AAB87413 represent the proteins they encode.

CC AAB87414-AAB87454 represent human secreted protein fragments. The genes

CC and their corresponding secreted proteins are useful for preventing,

CC treating or ameliorating medical conditions, e.g., by protein or gene

CC therapy. Pathological conditions can be diagnosed by determining the

CC amount of the new protein in a sample or by determining the presence of

CC mutations in the new genes. Specific uses are described for each of the

CC 52 genes, based on the tissues in which they are most highly expressed,

CC and include developing products for the diagnosis or treatment of

CC proliferative disorders, cancer, tumours, foetal and developmental

CC abnormalities, haematopoietic disorders, diseases of the immune system,

CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,

CC allergies, neurological disorders (e.g., Alzheimer's disease,

CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,

CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,

CC cardiovascular disorders, angiogenic disorders, kidney disorders,

CC gastrointestinal disorders, pregnancy-related disorders, endocrine

CC disorders, and infections. The proteins can also be used to aid wound

CC healing and epithelial cell proliferation, to prevent skin aging due to

CC sunburn, to maintain organs before transplantation, for supporting cell

CC culture of primary tissues, to regenerate tissues, to identify their

CC cognate ligands or binding partners, and in chemotaxis, and can be used

CC as a food additive or preservative to modify storage properties.

CC Antibodies specific for a protein of the invention can be used in

CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunosorbent assay (ELISA). The present sequence represents a human
CC secreted protein fragment referred to in the disclosure of the invention.
XX
SQ Sequence 257 AA;

Query Match 53.2%; Score 250; DB 22; Length 257;
Best Local Similarity 100.0%; Pred. No. 1.7e-232;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27 MDFTQHLGVCCLVALLISVGLLSVAACWFLPSIIIAAASWIITCVLLCCSKHARCFILLVF 86
Db 1 mdfiqhlgvccvlavllsvglsvaacwflpsiaaaswiitcvllccskharcfillvlf 60
Qy 87 LSCGLREGRNALIAAGTGIVILGHVENIFHFKGLLDGMCNLRKAKSFHIFPLKKYIE 146
Db 61 lscglregrnaliaaagtgvilghvenifhfkglldgmcnlraksfshfpllkkyie 120
Qy 147 AIQWYIGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 206
Db 121 aiqwiylglatplsvfddlvswngtlavslfshvleaqldnsgkevlsvlyqmatteev 180
Qy 207 LSSLGOKLLAFAGLSVLVLLGTGLFMKRFGLPGCGWKYENIYITRQFVQFDERERHQORPCV 266
Db 181 lsslgokllafaglsvlvllgtglfmkrflgpcgwkkyenyitrqfvqfdererhqgrpcv 240
Qy 267 LPLNKEERRK 276
Db 241 lplnkeerrk 250

RESULT 7
AAB34797
ID AAB34797 standard; Protein; 292 AA.
XX
AC AAB34797;
XX
DT 26-JAN-2001 (first entry)
XX
DE Human secreted protein sequence encoded by gene 25 SEQ ID NO:85.
XX
KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
KW neutropic; neuroprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antitumor; vulnery; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiant; gene therapy;
KW cancer; immune disorder; cardiovascular disorder; wound healing;
KW neurological disease; infectious disease; chromosome identification.

XX Homo sapiens.
OS
XX W0200058356-A1.
XX
XX 05-OCT-2000.
XX
XX 22-MAR-2000; 2000WO-US07535.
XX
XX 26-MAR-1999; 99US-0126511.
XX 17-DEC-1999; 99US-0172413.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Komatsoulis G;
XX
XX WPI; 2000-594639/56.
XX N-PSDB; AAC59990.
XX
XX Fifty nucleic acid molecules encoding human secreted proteins, useful
PT in the prevention, treatment and diagnosis of cancer, immune disorders,
PT cardiovascular disorders and neurological diseases -
XX
XX Claim 1; Page 385-386; 425pp; English.

XX The polynucleotide sequences given in AAC59966 to AAC60015 encode the
CC human secreted proteins given in AAB34773 to AAB34822. AAB34823 to
CC AAB34852 represent human secreted polypeptide sequences and proteins
CC homologous to them, which are given in the exemplification of the present
CC invention. Human secreted proteins have activities based on the tissues
CC and cells the genes are expressed in. Examples of activities include:
CC cytostatic; immunosuppressive; neutropic; neuroprotective; antiviral;
CC antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
CC vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic; and
CC cardiant. The polynucleotides and polypeptides are useful for
CC preventing, treating or ameliorating a medical condition in e.g. humans,
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. The polynucleotides are
CC useful for chromosome identification. They are also useful as probes for
CC diagnosing a disorder related to the female reproductive system,
CC particularly breast and/or ovary cancer. They are also useful in the gene
CC therapy of breast and ovarian cancer. Nucleic acids, protein, antibodies,
CC agonists and antagonists from the present invention are useful in the
CC diagnosis, treatment and prevention of cancer, immune disorders,
CC cardiovascular disorders, wound healing, neurological diseases and
CC infectious disease. AAC5957 to AAC5965 and AAB34772 represents sequence
CC used in the exemplification of the present invention.

XX Sequence 292 AA;

Query Match 50.9%; Score 239; DB 21; Length 292;
Best Local Similarity 100.0%; Pred. No. 7.4e-222;
Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27 MDFTQHLGVCCLVALLISVGLLSVAACWFLPSIIIAAASWIITCVLLCCSKHARCFILLVF 86
Db 1 mdfiqhlgvccvlavllsvglsvaacwflpsiaaaswiitcvllccskharcfillvlf 60
Qy 87 LSCGLREGRNALIAAGTGIVILGHVENIFHFKGLLDGMCNLRKAKSFHIFPLKKYIE 146
Db 61 lscglregrnaliaaagtgvilghvenifhfkglldgmcnlraksfshfpllkkyie 120
Qy 147 AIQWYIGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 206
Db 121 aiqwiylglatplsvfddlvswngtlavslfshvleaqldnsgkevlsvlyqmatteev 180
Qy 207 LSSLGOKLLAFAGLSVLVLLGTGLFMKRFGLPGCGWKYENIYITRQFVQFDERERHQORPC 265
Db 181 lsslgokllafaglsvlvllgtglfmkrflgpcgwkkyenyitrqfvqfdererhqgrpc 239

RESULT 8
AAB87401
ID AAB87401 standard; Protein; 291 AA.
XX
AC AAB87401;
XX
DT 22-MAY-2001 (first entry)
XX
XX Human gene 16 encoded secreted protein HMADJ74, SEQ ID NO:142.
DE
XX Human; secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angiogenic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
KW endocrine disorder; infection; wound healing; vulnery;
KW cell culture; chemotaxis; food additive;
KW binding partner identification.
XX
OS Homo sapiens.
XX


```

RESULT 10
AAR54005
ID AAR54005 standard; peptide; 45 AA.
XX
AC AAR54005;
XX
DT 03-NOV-1994 (first entry)
XX
DE Characteristic protein anchor motif.
XX
KW Conjugate vaccine; Streptococcus infection; group B; polysaccharide;
KW C protein alpha antigen; neonatal sepsis; meningitis; bca.
XX
OS Streptococcus.
XX
PN WO9410317-A.
XX
PD 11-MAY-1994.
XX
PF 02-NOV-1993; 93WO-US10506.
XX
PR 02-NOV-1992; 92US-0968866.
XX
PA (BGHM ) BRIGHAM & WOMENS HOSPITAL.
PA (GEOH ) GEN HOSPITAL CORP.
XX
PI Ausubel FM, Kasper DL, Madoff LC, Michel JL;
XX WPI; 1994-167472/20.
XX
PT New conjugate vaccine protects against group B Streptococcus
PT infection - comprises gp. B Streptococcus polysaccharide
PT conjugated to C protein alpha antigen deriv., useful against e.g.
PT neonatal sepsis and meningitis
XX
PS Disclosure; Page 81; 103pp; English.
XX
CC The sequence of the group B Streptococcus protein C (bca) was used
CC in a search to find homologous proteins. A class of Gram-positive
CC surface proteins with a common membrane anchor motif were found to
CC be homologous. The amino acid compen. at the C terminus of the
CC protein is characteristic of a protein membrane anchor.
CC See also AAR53996-4012.
XX
SQ Sequence 45 AA;

Query Match 1.5%; Score 7; DB 15; Length 45;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 SVGLLSV 49
Db 33 svglslv 39

RESULT 11
AAY25696
ID AAY25696 standard; peptide; 54 AA.
XX
AC AAY25696;
XX
DT 30-SEP-1999 (first entry)
XX
DE Barley germin N-terminal leader peptide.
XX
KW Nectarin; secretion; signal; plant; nectar; treatment; disease;
KW oxalate deposition; urolithiasis; kidney stone; barley; germin;
XX
OS Hordeum vulgare.
XX
PN US5939288-A.
XX

```

```

PD 17-AUG-1999.
XX
PF 07-JUN-1995; 95US-0478259.
XX
PR 07-JUN-1995; 95US-0478259.
XX
PA (IOWA ) UNIV IOWA STATE RES FOUND INC.
XX Thornburg R;
XX
PI WPI; 1999-468403/39.
XX
PT Expressing protein into nectar using a recombinant vector encoding
PT nectarins
XX
PS Example 1; Fig 2; 16pp; English.
XX
CC This invention describes a novel method for expressing protein into
CC nectar which comprises introducing a nucleic acid construct encoding a
CC Nicotiana nectarin signal peptide into a plant cell, where the signal
CC peptide facilitates secretion of the protein into nectar. The vectors and
CC methods are used to produce nectarins which may then be used for treating
CC diseases or conditions related to deposition of oxalate e.g. urolithiasis
CC (kidney stones). This sequence represents the leader sequence of a barley
CC germin protein which is used to illustrate the method of the invention.
XX
SQ Sequence 54 AA;

Query Match 1.5%; Score 7; DB 20; Length 54;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 TLAVSLF 176
Db 6 tlavslf 12

RESULT 12
AAY70219
ID AAY70219 standard; Protein; 57 AA.
XX
AC AAY70219;
XX
DT 06-JUN-2000 (first entry)
XX
DE Mouse Interleukin-1 epsilon short splice variant.
XX
KW Interleukin-1 epsilon; IL-1; chromosome 2q; cytokine; antiinflammatory;
KW immunosuppressant; antirheumatic; antiarthritic; antipsoriatic;
KW gene mapping; immune system; treatment; inflammatory disease;
KW autoimmune disease; rheumatoid arthritis; inflammatory bowel disease;
KW psoriasis; mouse; splice variant.
XX
OS Mus sp.
XX
PN WO200011174-A1.
XX
PD 02-MAR-2000.
XX
PF - 20-AUG-1999; 99WO-US18771.
XX
PR 21-AUG-1998; 98US-0097413.
PR 31-AUG-1998; 98US-0098595.
PR 11-SEP-1998; 98US-0099974.
XX
PA (IMMV ) IMMUNEX CORP.
XX
PI Sims JE, Smith DE;
XX
DR WPI; 2000-237653/20.
DR N-PSDB; AAZ51249.
XX

```


PT Nucleotide sequences encoding human interleukin -1 epsilon, useful to
PT treat inflammatory and immune system-related diseases such as
PT rheumatoid arthritis and inflammatory bowel disease -
XX
XX Disclosure; Page 9; 76pp; English.
XX
XX The present protein sequence is that of mouse Interleukin-1 (IL-1)
XX epsilon short splice variant. Mouse IL-1 epsilon DNA can be used as
XX probe to identify human IL-1 epsilon by screening human genomic library.
XX Human IL-1 epsilon gene is mapped to chromosome 2q and is mainly
XX expressed in spleen, lymph node, thymus, tonsil and leucocyte tissues.
XX IL-1 epsilon is a cytokine, with antiinflammatory, immunosuppressant,
XX antirheumatic, antiarthritic and antipsoriatic activity. It can be used
XX in the treatment of inflammatory or autoimmune diseases such as
XX rheumatoid arthritis, inflammatory bowel disease and psoriasis. The DNA
XX sequence can be used in chromosome identification, gene mapping and
XX study of immune system.
XX Sequence 57 AA;
SQ
Query Match 1.5%; Score 7; DB 21; Length 57;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 325 SKQFQSL 331
DB 24 skqfql 30
|||||
RESULT 13
AAAY87310
ID AAY87310 standard; Protein: 75 AA.
XX
XX AAY87310;
XX
XX 11-MAY-2000 (first entry)
XX
XX Human signal peptide containing protein HSPP-87 SEQ ID NO:87.
DE
XX Human; signal peptide-containing protein; HSPP; diagnosis; cancer;
XX inflammation; cardiovascular disease; anticancer; anti-inflammatory;
XX antimicrobial; neutropic; neuroprotective; cardiovascular; hepatotropic;
XX antiaesthetic; gene therapy; cell proliferation; neurological disorder;
XX reproductive disorder; developmental disorder; arteriosclerosis;
XX cirrhosis; psoriasis; acquired immune deficiency syndrome; anaemia;
XX asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;
XX Parkinson's disease; Huntington's disease; ovulatory defect;
XX muscular dystrophy.
XX
XX Homo sapiens.
XX
XX WO200000610-A2.
XX
XX 06-JAN-2000.
XX
XX 25-JUN-1999; 99WO-US14484.
XX
XX 26-JUN-1998; 98US-0090762.
XX
XX 31-JUL-1998; 98US-0094983.
XX
XX 01-OCT-1998; 98US-0102686.
XX
XX 11-DEC-1998; 98US-0112129.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Lal P, Tang YT, Gorgone GA, Corley NC, Guegler KJ, Baughn MR;
XX Akerbloom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JL;
XX Bandman O;
XX
XX WPI: 2000-160673/14.
XX
XX N-PSDB; AA98195.
XX
XX New human signal peptide-containing proteins useful in treatment,
PT

PT prevention and diagnosis of e.g. cancer, inflammation and
PT cardiovascular disease
XX
XX Claim 1; Page 216; 327pp; English.
XX
XX AA98109 to AA298242 encode AAY87224 to AAY87357 which represent the
XX human signal peptide-containing proteins HSPP-1 to HSPP-134. HSPPs have
XX anticancer, anti-inflammatory, antimicrobial, neutropic, hepatotropic,
XX neuroprotective, cardiovascular and antiaesthetic activities, and can
XX be used in gene therapy. HSPPs can be used to treat or prevent disorders
XX associated with decreased activity or function of HSPP. Antagonists of
XX HSPP are used to treat or prevent disorders associated with increased
XX activity or function of HSPP. Such diseases include cell proliferation
XX (including cancer), inflammation, cardiovascular, neurological,
XX reproductive or developmental disorders, (e.g. arteriosclerosis,
XX cirrhosis, psoriasis, acquired immune deficiency syndrome, anaemia,
XX asthma, Crohn's disease, microbial or other infections, congestive or
XX ischaemic heart disease, Alzheimer's, Parkinson's or Huntington's
XX diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSPP
XX nucleic acids can be used for the recombinant production of HSPP, for
XX detecting HSPP in standard hybridisation and amplification assays (for
XX diagnosis and monitoring), in gene therapy, as antisense,
XX triplex-forming or ribozyme therapeutics, for detecting related sequences
XX or genetic variations, and for chromosomal mapping. HSPP are also used to
XX raise specific antibodies (Ab) and to screen for agonists and
XX antagonists (potential therapeutic agents). Ab are used to diagnose, or
XX monitor, HSPP-related diseases (in usual immunoassays), as therapeutic
XX antagonists, in competitive drug screens, and for purification of HSPP
XX from natural sources.
XX
XX Sequence 75 AA;
SQ
Query Match 1.5%; Score 7; DB 21; Length 75;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 38 LVALISV 44
DB 8 lvalisv 14
|||||
RESULT 14
AAG55589
ID AAG55589 standard; Protein: 114 AA.
XX
XX AAG55589;
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 71305.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX
XX 05-MAR-1999; 99US-0123180.
XX
XX 09-MAR-1999; 99US-0123548.
XX
XX 23-MAR-1999; 99US-0125788.
XX
XX 25-MAR-1999; 99US-0126264.
XX
XX 29-MAR-1999; 99US-0126785.
XX
XX 01-APR-1999; 99US-0127462.
XX
XX 06-APR-1999; 99US-0128234.
XX
XX 08-APR-1999; 99US-0128714.
XX

PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.5%; Score 7; DB 21; Length 114;

Best Local Similarity 100.0%; Pred. No. 95;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 ALIAAGT 103

|||||

Db 70 allaagt 76

RESULT 15

AAG56533

ID AAG56533 standard; Protein; 114 AA.

XX AAG56533;

XX 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 72679.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

XX Arabidopsis thaliana.

OS EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145213.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 15-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156558.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.

PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162144.

Query Match 1.5%; Score 7; DB 21; Length 114;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 ALIAAGT 103
Db 70 aliaagt 76
|||||||

RESULT 16
AAG55588
ID AAG55588 standard; Protein; 121 AA.
XX
AC AAG55588;
XX
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 71304.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-01343768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-01351124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-014287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.5%; Score 7; DB 21; Length 121;
Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 ALIAAGT 103
|||||
Db 77 aliaagt 83

RESULT 17

AAG56532
ID AAG56532 standard; Protein; 122 AA.

XX AC AAG56532;

XX DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 72678.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126264.
PR 23-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.

PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 1.5%; Score 7; DB 21; Length 122;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 ALIAAGT 103
 |||||
 Db 78 aliaagt 84

RESULT 18
 AAY37783
 ID AAY37783 standard; Protein; 127 AA.
 XX
 AC AAY37783;
 XX
 DT 07-OCT-1999 (first entry)
 XX
 DE Protein involved in intermediate metabolism of nucleic acids.
 XX
 KW Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
 KW paratrachoma; inclusion conjunctivitis; genital disease; perihhepatitis;
 KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
 KW bartholinitis; pneumopathy; venereal lymphogranulomatosis.
 XX
 OS Chlamydia trachomatis.
 XX
 PN WO9928475-A2.
 XX
 PD 10-JUN-1999.
 XX
 PF 27-NOV-1998; 98WO-IB01939.
 XX
 PR 04-NOV-1998; 98US-0107077.
 PR 28-NOV-1997; 97FR-0015041.
 PR 17-DEC-1997; 97FR-0016034.
 XX
 PA (GEST) GENSET.
 XX
 PI Griffais R;
 XX
 WP1; 1999-371125/31.
 XX
 PT Genome sequence of Chlamydia trachomatis
 XX
 PS Disclosure; Page 1368; 1755pp; English.
 XX
 CC AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
 of Chlamydia trachomatis (see AA201425). The polypeptides can be used as
 vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
 can also be used to control growth of the microorganism. Chlamydia
 trachomatis is responsible for a large number of diseases, e.g. eye
 diseases such as conventional trachoma, nonendemic trachoma,
 paratrachoma, and inclusion conjunctivitis; genital diseases such as
 nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
 CC perihhepatitis, bartholinitis; pneumopathy in breast feeding infants;
 CC and venereal lymphogranulomatosis. The polypeptides of the invention
 CC may be of use in treating these diseases.
 XX
 SQ Sequence 127 AA;

Query Match 1.5%; Score 7; DB 20; Length 127;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 207 LSSLGQK 213
 |||||
 Db 14 lsslqk 20

PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 08-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 20-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159299.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 1.5%; Score 7; DB 21; Length 152;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 424 LLKRSK 430
 Db 139 llkrsk 145

RESULT 21

AA25909
ID AA25909 standard; Protein; 157 AA.

XX AC
XX AA25909;

XX DT 16-OCT-2001 (first entry)

XX DE Human protein sequence SEQ ID NO:1424.

XX KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
XX KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
XX KW antibacterial; endocrine; cardiant; central nervous system; virucide;
XX KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
XX KW antiaggregant; haemostatic; vulnery; antilulcer; osteopathic; eczema;
XX KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
XX KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
XX KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
XX KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
XX KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
XX KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
XX KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
XX KW allergic rhinitis; diabetes; multiple sclerosis; depression;
XX KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
XX KW neurological disorder.

XX OS Homo sapiens.

XX PN WO200153455-A2.

XX PD 26-JUL-2001.

XX PF 22-DEC-2000; 2000WO-US35017.

XX PR 23-DEC-1999; 99US-0471275.

XX PR 21-JAN-2000; 2000US-0488725.

XX PR 25-APR-2000; 2000US-0552317.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-457603/49.

XX DR N-PSDB; AAH99850.

XX PT Isolated human polynucleotides encoding polypeptides, useful for the

XX PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -

XX PS Claim 20; Page 289; 1217pp; English.

XX CC AAH99166 to AAH99904 encode the human proteins given in AA25225 to
XX CC AA25963. The proteins can have activities based on the tissues and
XX CC cells they are expressed in, such as: antiinflammatory; antirheumatic;
XX CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
XX CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
XX CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;
XX CC antilulcer; osteopathic; dermatological; antiallergic; antiasthmatic;
XX CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
XX CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
XX CC encoding them can be used in gene therapy, antisense therapy and vaccine
XX CC production. The proteins and polynucleotides are useful for screening for
XX CC agonists or antagonists of a protein and for the treatment and diagnosis
XX CC of disorders associated with the activity of a protein e.g. inflammation,
XX CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
XX CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
XX CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
XX CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
XX CC osteoporosis, severe combined immunodeficiency, eczema, allergic
XX CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
XX CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
XX CC neurological disorders.

XX Sequence 157 AA;

Query Match 1.5%; Score 7; DB 22; Length 157;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 LPLNREE 273

Db 118 lplnkee 124

RESULT 22

AA75031

ID AA75031 standard; Protein; 162 AA.

XX AC AA75031;

XX DT 21-MAR-2000 (first entry)

XX DE Neisseria meningitidis ORF 531 protein sequence SEQ ID NO:1536.

XX KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
XX KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
XX KW antibacterial; gene therapy.

XX OS Neisseria meningitidis.

XX PN WO9957280-A2.

XX PD 11-NOV-1999.

XX PF 30-APR-1999; 99WO-US09346.

XX PR 01-MAY-1998; 98US-0083758.

XX PR 31-JUL-1998; 98US-0094869.

XX PR 02-SEP-1998; 98US-0098994.

XX PR 02-SEP-1998; 98US-0099062.

XX PR 09-OCT-1998; 98US-0103749.

XX PR 09-OCT-1998; 98US-0103794.

XX PR 09-OCT-1998; 98US-0103796.

XX PR 25-FEB-1999; 99US-0121528.

XX PA (CHIR) CHIRON CORP.

XX PI (GENO-) INST GENOMIC RES.

XX PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;

XX PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarcelli M;

XX PI Tettelin H, Venter JC;

XX WPI; 2000-062150/05.

XX DR N-PSDB; AA253793.

XX PT Novel Neisserial polypeptides predicted to be useful antigens for
XX PT vaccines and diagnostics -
XX PS Claim 2; Page 802; 1453pp; English.
XX CC AA253015 to AA254536, AA254577 to AA254615, and AA254253 to AA254941
XX CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
XX CC and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
XX CC PCR primers used in the exemplification of the present invention. The
XX CC polypeptides, the polynucleotides, antibodies and compositions of
XX CC the invention can be used as vaccines as diagnostic reagents, and as
XX CC immunogenic compositions. The polypeptides can be used in the
XX CC manufacture of medicaments for treating or preventing infection due to
XX CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
XX CC presence of Neisseria bacteria, or to raise antibodies. They may also
XX CC be used to screen for agonists or antagonists, which may themselves
XX CC have use as antibacterial agents. The polynucleotides of the invention
XX CC may also be used in gene therapy protocols.

XX Sequence 162 AA;

Query Match 1.5%; Score 7; DB 21; Length 162;

Best Local Similarity .100.08; Pred. No. 1.3e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 ALIAAGT 103

Db 11 aliaagt 17

RESULT 23

AAG09373

ID AAG09373 standard; Protein; 163 AA.

XX AAG09373;

17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 7282.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 29-MAR-1999; 99US-0126264.

PR 01-APR-1999; 99US-0126785.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 19-MAY-1999; 99US-0134768.

PR 20-MAY-1999; 99US-0134941.

PR 21-MAY-1999; 99US-0135124.

PR 24-MAY-1999; 99US-0135353.

PR 25-MAY-1999; 99US-0135623.

PR 27-MAY-1999; 99US-0136021.

PR 28-MAY-1999; 99US-0136392.

PR 01-JUN-1999; 99US-0136782.

PR 03-JUN-1999; 99US-0137222.

PR 04-JUN-1999; 99US-0137528.

PR 07-JUN-1999; 99US-0137502.

PR 08-JUN-1999; 99US-0137724.

PR 10-JUN-1999; 99US-0138094.

PR 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 17-JUN-1999; 99US-0139453.

PR 18-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150366.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.5%; Score 7; DB 21; Length 163;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 LKRSK 430

|||||||

Db 150 llkrsk 156
RESULT 24
AAG16160
ID AAG16160 standard; Protein; 164 AA.
XX AC AAG16160;
XX XX
DT 17-OCT-2000 (first entry)
XX XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 16695.
XX XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX XX
OS Arabidopsis thaliana.
XX XX
PN EP1033405-A2.
XX XX
PD 06-SEP-2000.
XX XX
PF 25-FEB-2000; 2000EP-0301439.
XX XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139753.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0143977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 22-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 01-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 1.5%; Score 7; DB 21; Length 164;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 397 LMQLKIL 403
 |||||
 Db 19 lmqlkil 25

RESULT 25
 AAG16159
 ID AAG16159 standard; Protein; 187 AA.
 XX

AC AAG16159;
XX
DT 17-OCT-2000 (first entry)
DE
DE Arabidopsis thaliana protein fragment SEQ ID NO: 16694.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 03-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 25-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.

PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.5%; Score 7; DB 21; Length 187;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 397 LMOLKIL 403
Db 42 lmqkll 48
|||||

RESULT 26

AAB54077
ID AAB54077 standard; Protein; 192 AA.

XX AAB54077;

XX 09-MAR-2001 (first entry)

DE Human pancreatic cancer antigen protein sequence SEQ ID NO:529.

XX Human; pancreas; pancreatic cancer; pancreatic cancer antigen;

KW detection; diagnosis; identification; cytostatic; neuroprotective;
KW nontropic; immunomodulatory; relaxant; contraceptive; gynaecological;
KW antiinflammatory; cardiant; gene therapy; chromosome mapping;
KW linkage analysis; tissue identification; tissue typing; forensic;
KW neural; immune system; muscular; reproductive; gastrointestinal;
KW pulmonary; cardiovascular; renal; proliferative.
XX Homo sapiens.
OS
XX WO200055320-A1.
PN
XX 21-SEP-2000.
PD
XX 08-MAR-2000; 2000WO-US05989.
PF
XX 12-MAR-1999; 99US-0124270.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Ruben SM;
PI
XX WPI; 2000-579444/54.
DR
XX N-PSDB; AAC98842.
DR
XX New nucleic acid that is a pancreatic cancer antigen for preventing,
PT treating, or ameliorating a medical condition, particular pancreatic
PT cancer, or for use in assays for diagnosing a pathological condition -
XX
PS Claim 11; Page 966; 1379pp; English.
XX
CC AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC proteins, called pancreatic cancer antigens, given in AAB54008 to
CC AAB54466. The human pancreatic cancer antigens have cytostatic,
CC neuroprotective, nontropic, immunomodulatory, relaxant, contraceptive,
CC gynaecological, cardiant and antiinflammatory activities, and can be used
CC in gene therapy. The polynucleotide and proteins can be used for
CC preventing, treating, or ameliorating a medical condition or in assays
CC for diagnosing a pathological condition or a susceptibility to one in a
CC subject. Binding partners to the proteins and the activity of the
CC proteins can be identified. The pancreatic cancer antigens can be used to
CC detect, treat or prevent pancreatic disorders, especially cancer.
CC Agonists and antagonists to the antigens can be screened for. The
CC pancreatic cancer antigen polynucleotides can be used to design nucleic
CC acid hybridisation probes that can be used in chromosome mapping, linkage
CC analysis, tissue identification and/or typing and a variety of forensic
CC and diagnostic methods. The proteins can be used to generate antibodies
CC which are used to purify, detect and target the polypeptides, including
CC both in vivo and in vitro diagnostic and therapeutic methods. The
CC proteins can be used to treat or prevent neural, immune system, muscular,
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 192 AA;

Query Match 1.5%; Score 7; DB 21; Length 192;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 LSLVLLG 226
Db 4 lslvllg 10
|||||

RESULT 27

AAY18084

ID AAY18084 standard; Protein; 198 AA.

XX AAY18084;

XX 06-AUG-1999 (first entry)

DT

XX

DE Histamine binding protein AV-HBP.

KW Histamine binding protein; serotonin binding compound; inflammation;
 KW gastric acid secretion; allergy; type I hypersensitivity reaction;
 KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;
 KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;
 KW respiratory disease; coronary heart disease; cellular growth regulator;
 KW tissue repair; blood-sucking ectoparasite; therapy.

OS Amblyomma variegatum.

PN W09927104-A1.

XX 03-JUN-1999.

XX 26-NOV-1998; 98WO-GB03530.

XX 26-JUN-1998; 98GB-0013917.

PR 26-NOV-1997; 97GB-0025046.

XX (OXFO-) OXFORD VACS LTD.

XX Nuttall PA, Paesen GC;
 DR WPI; 1999-357841/30.
 DR N-PSDB; AAX76969.

XX Histamine and serotonin binding compounds useful for the treatment
 of allergies

XX Claim 12; Fig 6; 84pp; English.

XX This sequence is an example of a histamine or serotonin binding
 compound (A), of the invention. The compounds are useful for regulating
 the action of histamine and serotonin (in e.g. inflammation and gastric
 acid secretion), the detection, quantification and removal of histamine
 or serotonin (in animals, plants, cell cultures, food materials, or
 humans) and in the treatment of various diseases and allergies
 (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic
 rhinitis (hay fever), atopic dermatitis, insect bites and food and drug
 allergies, abnormal blood pressure, migraine, psychological disorders,
 respiratory disease, and coronary heart disease). Histamine may also be
 used to regulate cellular growth and tissue repair. The molecules may
 also be used as components of vaccines directed against blood-sucking
 ectoparasites.

XX Sequence 198 AA;

Query Match 1.5%; Score 7; DB 20; Length 198;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 DGMTCNL 129
 Db 136 dgmctcnl 142

RESULT 28
 AAY37507

ID AAY37507 standard; Protein; 200 AA.

XX AAY37507;

XX 07-OCT-1999 (first entry)

DE Protein which is specific to Chlamydia trachomatis.

XX Vaccine; eye disease; conjunctivitis; nonendemic trachoma;
 KW paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;
 KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
 KW bartholinitis; pneumonia; venereal lymphogranulomatosis.

XX

Chlamydia trachomatis.

W09928475-A2.

10-JUN-1999.

27-NOV-1998; 98WO-IB01939.

04-NOV-1998; 98US-0107077.

28-NOV-1997; 97FR-0015041.

17-DEC-1997; 97FR-0016034.

(GEST) GENSET.

Griffais R;

WPI; 1999-371125/31.

Genome sequence of Chlamydia trachomatis

Disclosure; Page 1183; 1755pp; English.

AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
 of Chlamydia trachomatis (See AA201425). The polypeptides can be used as
 vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
 can also be used to control growth of the microorganism. Chlamydia
 trachomatis is responsible for a large number of diseases, e.g. eye
 diseases such as conventional trachoma, nonendemic trachoma,
 paratrachoma, and inclusion conjunctivitis; genital diseases such as
 nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
 perihepatitis, bartholinitis; pneumonia in breast feeding infants;
 and venereal lymphogranulomatosis. The polypeptides of the invention
 may be of use in treating these diseases.

SQ Sequence 200 AA;

Query Match 1.5%; Score 7; DB 20; Length 200;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GLLSVAA 51

Db 64 gllsvaa 70

RESULT 29

AAE03298

ID AAE03298 standard; Protein; 217 AA.

XX AAE03298;

XX 10-AUG-2001 (first entry)

Human gene 7 encoded secreted protein HCRNC80, SEQ ID NO:72.

Human; secreted protein; proliferative disorder; cancer; tumour;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angioedema; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder;
 KW endocrine disorder; infection; wound healing; vulnerability;
 KW cell culture; chemotaxis; food additive; gene therapy;
 KW binding partner identification; chromosome 17.

OS Homo sapiens.

XX Key Location/Qualifiers

FH Peptide 1..20

FT /label= Signal_peptide

FT Protein 21..217
 TT /note= "Mature secreted protein"
 XX WO200134800-A1.
 XX 17-MAY-2001.
 XX
 XX 08-NOV-2000; 2000WO-US30674.
 XX
 XX 12-NOV-1999; 99US-0164750.
 XX 30-JUN-2000; 2000US-0215128.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Ruben SM, Komatsoulis GA, Ebner R, Fiscella M, Wei P;
 PI WPI; 2001-329085/34.
 XX N-PSDB; AAD07711.
 XX
 XX New nucleic acid molecules encoding human secreted proteins, used in
 PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
 PT Parkinson's diseases and cancers -
 XX
 XX Claim 11; Page 461; 530pp; English.
 XX
 XX AAD07705-AAD07759 represent cDNAs corresponding to 19 human secreted
 CC protein genes, and AAE03292-AAE03346 represent the proteins they encode.
 CC AAE03347-AAE03375 represent human secreted protein fragments or variants.
 CC The genes and their secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 19 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC proliferative disorders, cancer, tumours, foetal and developmental
 CC abnormalities, hematopoietic disorders, diseases of the immune system,
 CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
 CC allergies, neurological disorders (e.g., Alzheimer's disease,
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiotensin disorders, kidney disorders,
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine
 CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties.
 CC Antibodies specific for a protein of the invention can be used in
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein of the invention.
 XX
 XX Sequence 217 AA;
 SQ
 Query Match 1.5%; Score 7; DB 22; Length 217;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 214 LLAFAGL 220
 Db 170 llafagl 176
 |||||
 RESULT 30
 AAG90175
 ID AAG90175 standard; Protein; 222 AA.
 XX
 AC AAG90175;
 XX

DT 26-SEP-2001 (first entry)
 XX
 DE C glutamicum protein fragment SEQ ID NO: 3929.
 XX
 XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis.
 XX
 XX Corynebacterium glutamicum.
 OS
 XX EP1108790-A2.
 PN
 XX 20-JUN-2001.
 PD
 XX
 XX 18-DEC-2000; 2000EP-0127688.
 PF
 XX 16-DEC-1999; 99JP-0377484.
 PR 07-APR-2000; 2000JP-0159162.
 XX 03-AUG-2000; 2000JP-0280988.
 PR
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 PA
 XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX
 XX WPI; 2001-376931/40.
 DR N-PSDB; AAH65394.
 XX
 XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile or pattern of a gene and identifying homologous gene
 PT
 XX
 XX Claim 17; SEQ ID NO: 3929; 246pp + Sequence Listing; English.
 PS
 XX The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from Coryneform bacterium, and identifying a homologue of a gene derived
 CC from coryneform bacterium. Coryneform bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described
 CC in the exemplification of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.
 XX
 XX Sequence 222 AA;
 SQ
 Query Match 1.5%; Score 7; DB 22; Length 222;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 96 NALIAAG 102
 Db 145 naliaag 151
 |||||
 Search completed: January 31, 2002, 08:11:20
 Job time: 216 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2002, 08:03:09 ; Search time 28.55 Seconds
(without alignments)
370.457 Million cell updates/sec

Title: US-09-713-098-2
Perfect score: 470
Sequence: 1 MGITSTGDFLSLWEIYVS.....LPVLKMRKKOMDMSADSKS 470

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0
Searched: 212252 seqs, 22503292 residues

Word size : 0
Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : Issued Patents AA:*
1: /cgn2.6/ptodata/2/1aa/5A-COMB.pep:*
2: /cgn2.6/ptodata/2/1aa/5B-COMB.pep:*
3: /cgn2.6/ptodata/2/1aa/6A-COMB.pep:*
4: /cgn2.6/ptodata/2/1aa/6B-COMB.pep:*
5: /cgn2.6/ptodata/2/1aa/PCTUS-COMB.pep:*
6: /cgn2.6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	1.5	45	1 US-08-363-311-22	Sequence 22, Appl
2	7	1.5	45	2 US-08-463-288A-22	Sequence 22, Appl
3	7	1.5	45	2 US-08-470-445A-22	Sequence 22, Appl
4	7	1.5	45	2 US-08-462-679-22	Sequence 22, Appl
5	7	1.5	45	2 US-08-466-210A-22	Sequence 22, Appl
6	7	1.5	45	2 US-08-467-147A-22	Sequence 22, Appl
7	7	1.5	45	2 US-08-469-014-22	Sequence 22, Appl
8	7	1.5	45	5 PCT-US93-10506A-22	Sequence 22, Appl
9	7	1.5	45	5 PCT-US93-10506-22	Sequence 22, Appl
10	7	1.5	54	2 US-08-478-259-2	Sequence 22, Appl
11	7	1.5	235	1 US-07-869-933-34	Sequence 34, Appl
12	7	1.5	235	1 US-08-201-879A-5	Sequence 34, Appl
13	7	1.5	235	4 US-09-103-663-34	Sequence 34, Appl
14	7	1.5	278	5 PCT-US96-03916-60	Sequence 60, Appl
15	7	1.5	278	5 PCT-US96-03916-72	Sequence 72, Appl
16	7	1.5	355	4 US-08-818-112-79	Sequence 79, Appl
17	7	1.5	364	1 US-08-363-311-15	Sequence 15, Appl
18	7	1.5	364	2 US-08-463-288A-15	Sequence 15, Appl
19	7	1.5	364	2 US-08-470-445A-15	Sequence 15, Appl
20	7	1.5	364	2 US-08-462-679-15	Sequence 15, Appl
21	7	1.5	364	2 US-08-466-210A-15	Sequence 15, Appl
22	7	1.5	364	2 US-08-467-147A-15	Sequence 15, Appl
23	7	1.5	364	2 US-08-469-014-15	Sequence 15, Appl
24	7	1.5	364	5 PCT-US93-10506A-15	Sequence 15, Appl
25	7	1.5	364	5 PCT-US93-10506-15	Sequence 15, Appl
26	7	1.5	449	1 US-08-831-753-1	Sequence 1, Appl
27	7	1.5	1231	3 US-08-904-263A-4	Sequence 4, Appl

28	7	1.5	1865	1	US-08-588-985-2	Sequence 2, Appl
29	7	1.5	1865	1	US-08-971-988-2	Sequence 2, Appl
30	6	1.3	18	1	US-07-725-331-33	Sequence 33, Appl
31	6	1.3	18	1	US-07-725-331-44	Sequence 44, Appl
32	6	1.3	18	5	PCT-US91-05047-33	Sequence 33, Appl
33	6	1.3	18	5	PCT-US91-05047-44	Sequence 44, Appl
34	6	1.3	19	1	US-08-596-985-3	Sequence 3, Appl
35	6	1.3	19	3	US-08-792-832A-36	Sequence 36, Appl
36	6	1.3	24	2	US-08-902-516-9	Sequence 9, Appl
37	6	1.3	26	2	US-08-146-028-70	Sequence 70, Appl
38	6	1.3	26	4	US-08-723-425A-70	Sequence 70, Appl
39	6	1.3	26	4	US-09-112-206-70	Sequence 70, Appl
40	6	1.3	36	2	US-08-146-028-68	Sequence 68, Appl
41	6	1.3	36	2	US-08-146-028-126	Sequence 126, Appl
42	6	1.3	36	4	US-08-723-425A-68	Sequence 68, Appl
43	6	1.3	36	4	US-08-723-425A-126	Sequence 126, Appl
44	6	1.3	36	4	US-09-112-206-68	Sequence 68, Appl
45	6	1.3	36	4	US-09-112-206-126	Sequence 126, Appl
46	6	1.3	39	3	US-08-851-843A-26	Sequence 26, Appl
47	6	1.3	39	4	US-08-974-549A-216	Sequence 216, Appl
48	6	1.3	39	4	US-08-854-050-26	Sequence 26, Appl
49	6	1.3	40	3	US-08-792-832A-54	Sequence 54, Appl
50	6	1.3	43	1	US-08-197-792-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-363-311-22
; Sequence 22, Application US/08363311
; Patent No. 6548241
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Ausubel, Frederick M.
; TITLE OF INVENTION: Conjugate Vaccine Against Group B
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-2678
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/363,311
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/968,866
; FILING DATE: 02-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimdata, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 0609,3740004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 466-0800
; TELEFAX: (202) 833-8716
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
US-08-363-311-22

Query Match 1.5%; Score 7; DB 1; Length 45;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 SVGLLSV 49
| | | | | | | |
Db 33 SVGLLSV 39

RESULT 2

US-08-463-288A-22
; Sequence 22, Application US/08463288A
; Patent No. 5820660
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Kasper, Dennis L.
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine For Group B
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,288A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363,311
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
; APPLICATION NUMBER: US 07/408,036
; FILING DATE: 15-SEP-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugalsky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 0609.2370007
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
US-08-463-288A-22

Query Match 1.5%; Score 7; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 SVGLLSV 49
* | | | | | | | |
Db 33 SVGLLSV 39

RESULT 3

US-08-470-445A-22
; Sequence 22, Application US/08470445A
; Patent No. 5843444
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Kasper, Dennis L.
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine Against Group B
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,445A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363,311
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
; APPLICATION NUMBER: US 07/408,036
; FILING DATE: 15-SEP-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugalsky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 0609.237000A
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
US-08-470-445A-22

Query Match 1.5%; Score 7; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 SVGLLSV 49
| | | | | | | |
Db 33 SVGLLSV 39

RESULT 4

US-08-462-679-22
; Sequence 22, Application US/08462679
; Patent No. 5847081
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Kasper, Dennis L.
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine For Group B

```

; TITLE OF INVENTION: Streptococcus
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,679
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363,311
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/408,036
; FILING DATE: 15-SEP-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugalsky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 0609.2370008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
; US-08-462-679-22

Query Match 1.5%; Score 7; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 SVGLLSV 49
Db 33 SVGLLSV 39

RESULT 5
US-08-466-210A-22
; Sequence 22, Application US/08466210A
; Patent No. 5858362
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Kasper, Dennis L.
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine For Group B
; TITLE OF INVENTION: Streptococcus
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,147A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363,311
```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,210A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363,311
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/408,036
; FILING DATE: 15-SEP-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugalsky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 0609.237000B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
; US-08-466-210A-22

Query Match 1.5%; Score 7; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 SVGLLSV 49
Db 33 SVGLLSV 39

RESULT 6
US-08-467-147A-22
; Sequence 22, Application US/08467147A
; Patent No. 5908629
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Kasper, Dennis L.
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine For Group B
; TITLE OF INVENTION: Streptococcus
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,147A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363,311
```

; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA: US 07/968,866
; APPLICATION NUMBER: 02-NOV-1992
; FILING DATE: 02-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/408,036
; FILING DATE: 15-SEP-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugalsky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 0609.2370009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2500
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
; US-08-467-147A-22

Query Match 1.5%; Score 7; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 SVGLLSV 49
|||||||
Db 33 SVGLLSV 39

RESULT 7
US-08-469-014-22
; Sequence 22, Application US/08469014
; Patent No. 5968521
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Kasper, Dennis L.
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine Against Group B
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,014
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363,311
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/408,036
; FILING DATE: 15-SEP-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugalsky, Lawrence B.
; REGISTRATION NUMBER: 35,086

; REFERENCE/DOCKET NUMBER: 0609.2370006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
; US-08-469-014-22

Query Match 1.5%; Score 7; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 SVGLLSV 49
|||||||
Db 33 SVGLLSV 39

RESULT 8
PCT-US93-10506A-22
; Sequence 22, Application PC/TUS9310506A
; GENERAL INFORMATION:
; APPLICANT: THE GENERAL HOSPITAL CORPORATION
; APPLICANT: Fruit Street
; APPLICANT: Boston, Massachusetts 02114
; APPLICANT: United States of America
; APPLICANT: 75 Francis Street
; APPLICANT: Boston, Massachusetts 02115
; APPLICANT: United States of America
; TITLE OF INVENTION: Conjugate Vaccine Against Group B
; TITLE OF INVENTION: Streptococcus
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Ave., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10506A
; FILING DATE: 02-NOV-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 0609.237PC01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
; PCT-US93-10506A-22

Query Match 1.5%; Score 7; DB 5; Length 45;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 SVGLSV 49
|||||||
Db 33 SVGLSV 39

RESULT 9

PCT-US93-10506-22
; Sequence 22, Application PC/TUS9310506
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Kasper, Dennis
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine Against Group B
; TITLE OF INVENTION: Streptococcus
; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, N.W.; Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10506
; FILING DATE: 02-NOV-1993
; CLASSIFICATION:

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.

; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 0609.237PC01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSK

; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
PCT-US93-10506-22

Query Match 1.5%; Score 7; DB 5; Length 45;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 SVGLSV 49
|||||||
Db 33 SVGLSV 39

RESULT 10

US-08-478-259-2
; Sequence 2, Application US/08478259
; Patent No. 5939288
; GENERAL INFORMATION:

; APPLICANT: Thornburg, Robert
; TITLE OF INVENTION: PLANT SECRETORY SIGNAL PEPTIDES AND NECTARINS
; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Baker & Botts, L.L.P.
; STREET: 910 Louisiana
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77002-4995

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,259
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:

; ATTORNEY/AGENT INFORMATION:
; NAME: Turley, Patrick
; REGISTRATION NUMBER: 35723
; REFERENCE/DOCKET NUMBER: 020053.0102
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-229-1791
; TELEFAX: 713-229-1522

; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 54 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Hordium vulgare Germin CM72
US-08-478-259-2

Query Match 1.5%; Score 7; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 TLAVSLF 176
|||||||
Db 6 TLAVSLF 12

RESULT 11

US-07-869-933-34
; Sequence 34, Application US/07869933
; Patent No. 5770396
; GENERAL INFORMATION:

; APPLICANT: KINET, Jean-Pierre
; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
; TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
; TITLE OF INVENTION: IMMUNOGLOBULIN
; NUMBER OF SEQUENCES: 34

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/869,933
; FILING DATE: 19920416
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/154 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: mouse
; STRAIN: FcRI beta subunit
; US-07-869-933-34

Query Match 1.5%; Score 7; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 99 IAAGTGI 105
|||||
Db 131 IAAGTGI 137

RESULT 12

US-08-201-879A-5
; Sequence 5, Application US/08201879A
; Patent No. 5807988
; GENERAL INFORMATION:

; APPLICANT: KINET, Jean-Pierre
; APPLICANT: JOUVIN, Marie-Helene
; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
; TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
; TITLE OF INVENTION: IMMUNOGLOBULIN E
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/201,879A
; FILING DATE: 24-FEB-1994

CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/869,933

; FILING DATE: 16-APR-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US93/03419

; FILING DATE: 16-APR-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: BENT, Stephen A.

; REGISTRATION NUMBER: 29,768

; REFERENCE/DOCKET NUMBER: 40399/234/NIHD

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)672-5300

; TELEFAX: (202)672-5399

; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-201-879A-5

Query Match 1.5%; Score 7; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 99 IAAGTGI 105
|||||
Db 131 IAAGTGI 137

RESULT 13

US-09-103-663-34
; Sequence 34, Application US/09103663D
; Patent No. 6171803
; GENERAL INFORMATION:

; APPLICANT: Kinet et al.
; TITLE OF INVENTION: Isolation, characterization, and use of the human beta
; TITLE OF INVENTION: subunit of the high affinity receptor for
; TITLE OF INVENTION: immunoglobulin E.
; FILE REFERENCE: 50490

; CURRENT APPLICATION NUMBER: US/09/103,663D

; CURRENT FILING DATE: 1998-06-23

; EARLIER APPLICATION NUMBER: 07/869,933

; EARLIER FILING DATE: 1992-04-16

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 34

; LENGTH: 235

; TYPE: PRT

; ORGANISM: Mus sp.

US-09-103-663-34

Query Match

Best Local Similarity 100.0%; Score 7; DB 4; Length 235;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 99 IAAGTGI 105
|||||
Db 131 IAAGTGI 137

RESULT 14

PCT-US96-03916-60

; Sequence 60, Application PC/TUS9603916

; GENERAL INFORMATION:

; APPLICANT: Wild, Martha A.

; APPLICANT: Cochran, Mark D.

; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS

; TITLE OF INVENTION: AND USES THEREOF

; NUMBER OF SEQUENCES: 72

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.25

; CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/03916
FILING DATE: 23-MAR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,597
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39116-A
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 278 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-03916-60

Query Match 1.5%; Score 7; DB 5; Length 278;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 RNALIAA 101
|||||
Db 80 RNALIAA 86

RESULT 15
PCT-US96-03916-72
Sequence 72, Application PC/TUS9603916
GENERAL INFORMATION:
APPLICANT: Wild, Martha A.
APPLICANT: Cochran, Mark D.
TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/03916
FILING DATE: 23-MAR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,597
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39116-A
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 278 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-03916-72

Query Match 1.5%; Score 7; DB 5; Length 278;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 RNALIAA 101
|||||
Db 80 RNALIAA 86

RESULT 16
US-08-818-112-79
Sequence 79, Application US/08818112
Patent No. 6290969
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, David C.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Houghton, Raymond
APPLICANT: Vedvick, Thomas S.
APPLICANT: Twardzik, Daniel R.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS
NUMBER OF SEQUENCES: 153
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/818,112
FILING DATE: 13-MAR-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.411C6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-818-112-79

Query Match 1.5%; Score 7; DB 4; Length 355;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 AGTGIVI 107
|||||
Db 78 AGTGIVI 84

RESULT 17
US-08-363-311-15
Sequence 15, Application US/08363311
Patent No. 6548241
GENERAL INFORMATION:
APPLICANT: Michel, James L.

; APPLICANT: Ausubel, Frederick M.
; TITLE OF INVENTION: Conjugate Vaccine Against Group B
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-2678
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/363,311
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/968,866
; FILING DATE: 02-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 0609.3740004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 466-0800
; TELEFAX: (202) 833-8716
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 310
; OTHER INFORMATION: /note= "This feature indicates that
; the amino acid sequence from position 227 through
; 309 is inserted at position 310 and may repeat up to
; eight times (for a total of nine repeating copies of
; these sequences within the polypeptide)."
; OTHER INFORMATION:
US-08-363-311-15

Query Match 1.5%; Score 7; DB 1; Length 364;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 SVGLLSV 49
| | | | |
Db 352 SVGLLSV 358

RESULT 18
US-08-463-288A-15
; Sequence 15, Application US/08463288A
; Patent No. 5820860
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Kasper, Dennis L.
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine For Group B
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: Washington

; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,288A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363,311
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/408,036
; FILING DATE: 15-SEP-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugaisky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 0609.2370007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 310
; OTHER INFORMATION: /note= "This feature indicates that
; the amino acid sequence from position 227 through
; 309 is inserted at position 310 and may repeat up to
; eight times (for a total of nine repeating copies of
; these sequences within the polypeptide)."
; OTHER INFORMATION:
US-08-463-288A-15

Query Match 1.5%; Score 7; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 SVGLLSV 49
| | | | |
Db 352 SVGLLSV 358

RESULT 19
US-08-470-445A-15
; Sequence 15, Application US/08470445A
; Patent No. 5843444
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Kasper, Dennis L.
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine Against Group B
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA

```
;
; ZIP: 20005-3934
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,445A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363,311
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
; APPLICATION NUMBER: US 07/408,036
; FILING DATE: 15-SEP-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugaisky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 0609.237000A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 310
; OTHER INFORMATION: /note= "this feature indicates that
; the amino acid sequence from position 227 through
; 309 is inserted at position 310 and may repeat up to
; eight times (for a total of nine repeating copies of
; these sequences within the polypeptide)."
;
; US-08-470-445A-15
;
; Query Match 1.5%; Score 7; DB 2; Length 364;
; Best Local Similarity 100.0%; Pred. No. 1.2e+02;
; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 43 SVGLLSV 49
; | | | | |
; Db 352 SVGLLSV 358
;
; RESULT 20
; US-08-462-679-15
; Sequence 15, Application US/08462679
; Patent No. 5847081
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Kasper, Dennis L.
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine For Group B
; TITLE OF INVENTION: Streptococcus
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
;
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,679
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363,311
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
; APPLICATION NUMBER: US 07/408,036
; FILING DATE: 15-SEP-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugaisky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 0609.2370008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 310
; OTHER INFORMATION: /note= "this feature indicates that
; the amino acid sequence from position 227 through
; 309 is inserted at position 310 and may repeat up to
; eight times (for a total of nine repeating copies of
; these sequences within the polypeptide)."
;
; US-08-462-679-15
;
; Query Match 1.5%; Score 7; DB 2; Length 364;
; Best Local Similarity 100.0%; Pred. No. 1.2e+02;
; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 43 SVGLLSV 49
; | | | | |
; Db 352 SVGLLSV 358
;
; RESULT 21
; US-08-466-210A-15
; Sequence 15, Application US/08466210A
; Patent No. 5858362
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Kasper, Dennis L.
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine For Group B
; TITLE OF INVENTION: Streptococcus
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
/
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/466,210A
/ FILING DATE: 06-JUN-1995
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/363,311
/ FILING DATE: 22-DEC-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/968,866
/ FILING DATE: 02-NOV-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/408,036
/ FILING DATE: 15-SEP-1989
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bugalsky, Lawrence B.
/ REGISTRATION NUMBER: 35,086
/ REFERENCE/DOCKET NUMBER: 0609.237000B
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 371-2600
/ TELEFAX: (202) 371-2540
/ TELEX: 248636 SSK
/ INFORMATION FOR SEQ ID NO: 15:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 364 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 310
/ OTHER INFORMATION: /note= "This feature indicates that
/ the amino acid sequence from position 227 through
/ 309 is inserted at position 310 and may repeat up to
/ eight times (for a total of nine repeating copies of
/ these sequences within the polypeptide)."
/
/ US-08-466-210A-15

Query Match 1.5%: Score 7; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 SVGLLSV 49
Db 352 SVGLLSV 358
|||||

RESULT 22
US-08-467-147A-15
; Sequence 15, Application US/08467147A
; Patent No. 5908629
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Kasper, Dennis L.
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine For Group B
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/467,147A
/ FILING DATE: 06-JUN-1995
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/363,311
/ FILING DATE: 22-DEC-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/968,866
/ FILING DATE: 02-NOV-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/408,036
/ FILING DATE: 15-SEP-1989
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bugalsky, Lawrence B.
/ REGISTRATION NUMBER: 35,086
/ REFERENCE/DOCKET NUMBER: 0609.2370009
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 371-2600
/ TELEFAX: (202) 371-2540
/ TELEX: 248636 SSK
/ INFORMATION FOR SEQ ID NO: 15:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 364 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 310
/ OTHER INFORMATION: /note= "This feature indicates that
/ the amino acid sequence from position 227 through
/ 309 is inserted at position 310 and may repeat up to
/ eight times (for a total of nine repeating copies of
/ these sequences within the polypeptide)."
/
/ US-08-467-147A-15

Query Match 1.5%: Score 7; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 SVGLLSV 49
Db 352 SVGLLSV 358
|||||

RESULT 23
US-08-469-014-15
; Sequence 15, Application US/08469014
; Patent No. 5968521
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Kasper, Dennis L.
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine Against Group B
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,014
```

;
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363,311
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/408,036
; FILING DATE: 15-SEP-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugalsky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 0609.2370006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 310
; OTHER INFORMATION: /note= "This feature indicates that
; the amino acid sequence from position 227 through
; 309 is inserted at position 310 and may repeat up to
; eight times (for a total of nine repeating copies of
; these sequences within the polypeptide)."
; OTHER INFORMATION: these sequences within the polypeptide)."
; US-08-469-014-15

Query Match 1.5%; Score 7; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 SVGLLSV 49
Db 352 SVGLLSV 358

RESULT 24
PCT-US93-10506A-15
; Sequence 15, Application PC/TUS9310506A
; GENERAL INFORMATION:
; APPLICANT: THE GENERAL HOSPITAL CORPORATION
; APPLICANT: Fruit Street
; APPLICANT: Boston, Massachusetts 02114
; APPLICANT: United States of America
; APPLICANT: 75 Francis Street
; APPLICANT: Boston, Massachusetts 02115
; APPLICANT: United States of America
; TITLE OF INVENTION: Conjugate Vaccine Against Group B
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Ave., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10506A

;
; FILING DATE: 02-NOV-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 0609.237PC01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 310
; OTHER INFORMATION: /note= "This feature indicates that
; the amino acid sequence from position 227 through
; 309 is inserted at position 310 and may repeat up to
; eight times (for a total of nine repeating copies of
; these sequences within the polypeptide)."
; OTHER INFORMATION: these sequences within the polypeptide)."
; PCT-US93-10506A-15

Query Match 1.5%; Score 7; DB 5; Length 364;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 SVGLLSV 49
Db 352 SVGLLSV 358

RESULT 25
PCT-US93-10506-15
; Sequence 15, Application PC/TUS9310506
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Kasper, Dennis
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine Against Group B
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, N.W.; Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10506
; FILING DATE: 02-NOV-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 0609.237PC01

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
TELEX: 248636 SSK

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:
LENGTH: 364 amino acids
TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

FEATURE:

NAME/KEY: misc_feature

LOCATION: 310

OTHER INFORMATION: /note= "this feature indicates that

the amino acid sequence from position 227 through

309 is inserted at position 310 and may repeat up to

eight times (for a total of nine repeating copies of

these sequences within the polypeptide)."

PCT-US93-10506-15

Query Match

1.5%; Score 7; DB 5; Length 364;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 SVGLLSV 49

|||||||

Db 352 SVGLLSV 358

RESULT 26

US-08-831-753-1

Sequence 1, Application US/08831753

Patent No. 5798234

GENERAL INFORMATION:

APPLICANT: ENGEL, PAUL C

APPLICANT: RICE, DAVID

TITLE OF INVENTION: METHOD FOR THE DIRECTED MODIFICATION OF

OPERATING SYSTEM: ENZYMES, MODIFIED ENZYMES AND THEIR USE

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: PILLSBURY MADISON & SUTRO LLP

STREET: 1100 NEW YORK AVENUE, NW, NINTH FLOOR, EAST

CITY: WASHINGTON

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/831,753

FILING DATE: 01-APR-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: HOBBS, ANN S

REGISTRATION NUMBER: 36830

REFERENCE/DOCKET NUMBER: 21123/235858

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-861-3000

TELEFAX: 202-822-0944

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 449 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Clostridium symbiosum

US-08-831-753-1

Query Match

1.5%; Score 7; DB 1; Length 449;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 EVLSSLG 211

|||||||

Db 27 EVLSSLG 33

RESULT 27

US-08-904-263A-4

Sequence 4, Application US/08904263A

Patent No. 6015889

GENERAL INFORMATION:

APPLICANT: LINDAHL, GUNNAR

APPLICANT: STALHAMMAR-CARLEMALM,

APPLICANT: MARGARETHA

TITLE OF INVENTION: STENBERG, LARS

TITLE OF INVENTION: PROTEIN RIB, A CELL SURFACE PROTEIN THAT

CONFERES IMMUNITY TO MANY STRAINS OF THE GROUP B

TITLE OF INVENTION: STREPTOCOCCUS: PROCESS FOR PURIFICATION OF THE PROTEIN,

TITLE OF INVENTION: REAGENT KIT AND PHARMACEUTICAL COMPOSITION

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP

STREET: PO BOX 747

CITY: FALLS CHURCH

STATE: VA

COUNTRY: USA

ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/904,263A

FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: SVENSSON, LEONARD R

REGISTRATION NUMBER: 30,330

REFERENCE/DOCKET NUMBER: 552-119P

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 205-8000

TELEFAX: (703) 205-8050

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1231 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-904-263A-4

Query Match

1.5%; Score 7; DB 3; Length 1231;

Best Local Similarity 100.0%; Pred. No. 3.4e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 SVGLLSV 49

|||||||

Db 1219 SVGLLSV 1225

RESULT 28

US-08-588-985-2

Sequence 2, Application US/08588985

Patent No. 5777094

GENERAL INFORMATION:

```

; APPLICANT: Michiyuki MATSUDA et al.
; TITLE OF INVENTION: cDNA OF DOCK180 GENE AND DOCK180 PROTEIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/588,985
; FILING DATE: January 19, 1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1865 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE: spleen cell of homo sapiens
; US-08-588-985-2
;
; Query Match 1.5%; Score 7; DB 1; Length 1865;
; Best Local Similarity 100.0%; Pred. No. 4.9e+02;
; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
QY 420 LHAKLLK 426
Db 1242 LHAKLLK 1248
;
; RESULT 29
; US-08-971-988-2
; Sequence 2, Application US/08971988
; Patent No. 5786451
; GENERAL INFORMATION:
; APPLICANT: Michiyuki MATSUDA et al.
; TITLE OF INVENTION: cDNA OF DOCK180 GENE AND DOCK180 PROTEIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971,988
; FILING DATE: 17-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/588,985
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1865 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE: spleen cell of homo sapiens
; US-08-971-988-2
;
; Query Match 1.5%; Score 7; DB 1; Length 1865;
; Best Local Similarity 100.0%; Pred. No. 4.9e+02;
; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
QY 420 LHAKLLK 426
Db 1242 LHAKLLK 1248
;
; RESULT 30
; US-07-725-331-33
; Sequence 33, Application US/07725331
; Patent No. 5294605
; GENERAL INFORMATION:
; APPLICANT: Houghten, Richard
; APPLICANT: Blondelle, Sylvie
; TITLE OF INVENTION: Amphiphilic Peptide Compositions and
; TITLE OF INVENTION: Analogues Thereof
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldsmith, Sutker, Shore,
; STREET: 180 No. 5294605th Stetson
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/725,331
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/554,422
; FILING DATE: 19-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER: 421250-80
; TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: 3126165418
; TELEFAX: 3126165460
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; OTHER INFORMATION: C-terminal amide, acetylated at
; OTHER INFORMATION: N-terminus.
US-07-725-331-33

Query Match 1.3%; Score 6; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 423 KLLKKR 428
| | | | |
Db 6 KLLKKR 11

Search completed: January 31, 2002, 08:10:18
Job time: 429 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2002, 08:00:39 ; Search time 33.39 Seconds
(without alignments)
1072.238 Million cell updates/sec

Title: US-09-713-098-2
 Perfect score: 2448
 Sequence: 1 MGIIWTSCTDIFLSLWEIYVS.....LPVLKMITRKKQMDNASADKS 470

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 219241

```

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

```
Database :      PIR_68:*
1:  pir1:*
2:  pir2:*
3:  pir3:*
4:  pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	113.5	4.6	351	2	T15112	hypothetical prote	
2	102.5	4.2	407	2	E71665	bicyclomyacin resis	
3	102	4.2	391	2	H81265	probable transmemb	
4	101	4.1	645	2	T28867	hypothetical prote	
5	100.5	4.1	729	2	I52481	PEPT 2 - human	
6	100	4.1	997	2	A04812	Ca2+-transporting	
7	100	4.1	997	2	S04651	Ca2+-transporting	
8	100	4.1	997	2	B40812	Ca2+-transporting	
9	100	4.1	1042	2	S04652	Ca2+-transporting	
10	99.5	4.1	470	2	G86516	dicarboxylase tran	
11	99.5	4.1	690	2	G84638	hypothetical prote	
12	99	4.0	510	2	CA30594	Ca2+-transporting	
13	99	4.0	997	1	PWRBSC	Ca2+-transporting	
14	99	4.0	997	2	B31981	Ca2+-transporting	
15	99	4.0	997	2	B31982	Ca2+-transporting	
16	99	4.0	997	2	S23444	Ca2+-transporting	
17	99	4.0	1042	1	PWRBMC	Ca2+-transporting	
18	99	4.0	1042	2	CA31981	Ca2+-transporting	
19	99	4.0	1042	2	A33881	Ca2+-transporting	
20	99	4.0	1043	2	A31982	Ca2+-transporting	
21	97.5	4.0	1065	2	E69795	acriflavin resista	
22	97	4.0	614	2	A69845	Na+/H+ antipater	
23	97	4.0	626	2	D70178	PTS system, fructo	
24	96.5	3.9	732	2	T23505	hypothetical prote	
25	96	3.9	405	2	D64522	carboxynorspermi	
26	96	3.9	451	2	G71561	probable integral	
27	96	3.9	514	2	S56384	hypothetical 56.3k	
28	95	3.9	486	2	T71683	NADH dehydrogenase	
29	94	3.8	500	2	A86112	probable transport	

ALIGNMENTS

RESULT 1

T15112
hypothetical protein ZC132.7 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jan-2000
C:Accession: T15112
R:Bradshaw, H.; Devlin, K.
submitted to the EMBL data Library, July 1997
A:Description: The sequence of *C. elegans* cosmid ZC132.
A:Reference number: Z18294
A:Accession: T15112
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-351 <BRA>
A:Cross-references: EMBL:AF014939; NID:Q2275620; PID:Q2275626; PIDN:AAB63929.1; GSPDB:
A:Experimental source: strain Bristol N2; clone ZC132
C:Genetics:
A:Gene: CESP:ZC132.7
A:Map position: 5
A:Introns: 118/1; 181/3
C:Superfamily: *Caenorhabditis* hypothetical protein C49G7.2

[illegible]

Db 332 SQ 333

RESULT 2

E71665

bicyclomycin resistance protein (bcrl) RP603 - Rickettsia prowazekii

C:Species: Rickettsia prowazekii

C>Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000

C:Accession: E71665

R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sacheritz-Ponten, T.; Alsmark, U.

A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.

A:Reference number: A71630; MUID:99039499

A:Accession: E71665

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-407 <AND>

A:Cross-references: GB:AJ235272; GB:AJ235269; NID:g3861033; PIDN:CAA15047.1; PID:g386114

A:Experimental source: strain Madrid E

C:Genetics:

A:Gene: bcrl; RP603

C:Superfamily: bicyclomycin resistance protein

Query Match

Best Local Similarity 4.2%; Score 102.5; DB 2; Length 407;

Matches 79; Conservative 79; Mismatches 131; Indels 195; Gaps 22;

QY 58 IIAAASWIIITCVLLC-----CSKHARCFIL-----LVPL--SCGLRE- 93

Db 3 IIAKIPAWMLLCFLFSPPTTETTYTGLPSLTRCFSDGCTQMTSLYLFAGVAFSL 62

QY 94 -----GRNALAAGTGVILGHVENIFHFNKGLDGMTCNLRKAKSFSTHFPLKKVIE 146

Db 63 GLRSDIYGRPIVLLGLFIYVSSIIF-----SFNEMLMIARFIQ 105

QY 147 A-----IQWTYGLATPLSVF-----DDLVSNNOTLA 172

Db 106 AFGVSGSVIGQSWARDSYOGAELSYVAILSPWLLFPALGSYIGYVIEIYLSWHYVFI 165

QY 173 VSLFSPSHVLEAQLNDSKGEVLSVLYOMATTEVLSLQKLLAFAGLSVLLGTGFMK 232

Db 166 --FPLAGTI-----LLALYYQILPETNYIAFSSQSKYFEFNIITKDKMLWLY 213

QY 233 RF-LGPCGKXENIYITQFVQFDERERHQPCVLPNKEERKYVLIPTFWPT----- 286

Db 214 AFIIGAFNGIYGFIEAPFILIDQMR-----VLPFSYKGLAFLL 253

QY 287 -----PKERNKLGFLFPLILHLICIWLVFAA-----VDYLLYRLI 321

Db 254 SFASIFGFLGGYLIKRRQYVDKKVWSIGFT-----SLCGCILFAVDSFILEFILYSNV 308

QY 322 FSVSKQFQSLPGFEVH-----LKLHGKQGTQDIHDSFNISFEPN-C 365

Db 309 FAIAMIIFPMV---MIHMGHSLIAITRLVALEDYATVGTAGSIFGAIYVVIASVTYC 365

QY 366 IPKPKFLLSTWPLSVLILLYMLGSLLSILMQLKILVSASFYPSVERKRIQYLHAKLL 425

Db 366 VSK---IHGETISNLSLCLVL-----SISVI-----YI-CILLY 399

QY 426 KRRS 429

Db 400 KKS 403

RESULT 3

H81265

probable transmembrane transport protein Cj1684c [imported] - Campylobacter jejuni (stra

C:Species: Campylobacter jejuni

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 31-Mar-2000

C:Accession: H81265

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillin

C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Ba
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals
A:Reference number: A81250; MUID:20150912
A:Accession: H81265
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-391 <PAR>
A:Cross-references: GB:AL139079; GB:AL111168; NID:g6968971; PIDN:CAB73670.1; PID:g696
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: Cj1684c

Query Match 4.2%; Score 102; DB 2; Length 391;
Best Local Similarity 20.9%; Pred. No. 0.52;
Matches 96; Conservative 60; Mismatches 131; Indels 172; Gaps 23;

QY 82 ILLVFLSCGLREGRNALI-----AAGTGIV---ILGH-----VENTFHNPFGLL 122

Db 16 ILIVIALCLLFSPIAKILRLPLSATEIILGAIIAHFGFIGKSENFALLANVGFYILMFI 75

QY 123 DGMTCNLRA-----KFSIHFPPLKKYIEALQWYGLATPLSVFDDLVSWNOTLA 172

Db 76 AGMEVNLRAFPNMDKEIAKKSEFYIF-LIYALSSFTVWIFGLSVFVIIIPVMSVG---- 130

QY 173 VSLFSPSHVLEAQLNDSKGEV--LSVLYOMATTEVLSLQKLLAFAGLSVLLGTG-- 228

Db 131 -----LLSLLFKDCKEYCNLNIAMIVATLAEVISIV---LITIAG-AFIREGTGII 178

QY 229 -----LPMKRFGLGPC-----GWKXENIYITQFVQFDERERHQPCVLPNKE 272

Db 179 DVAQSILYLNIFLGLCLLGFKMLGVFWWYPOI----- 211

QY 273 ERRKYVIPTWPTPKERNKGLFELPILILHLICIWLVFAVDYLLYRLI-----FSV 324

Db 212 ---KVLMP--W-EDNKEKDI-RFCMAIFILIIIVAMVITKLEIVLGSFAGSFIATFFDH 264

QY 325 SKQFQS-----LPGFEVHLKLGKQGTQDIHDSSEFNISVFEPNCPKPKFLS 374

Db 265 KDLKLEKLTGTHGFLIPIFFIHI-----GSTFDLK-----MIL 298

QY 375 ETWVPLSVILLILYMLGSLLSILMQLKILVSASFYPSVERKRIQYLHAKLLKRSQPL- 433

Db 299 DYKIVLQAFLLMFVWVG-----LRILCASVEL-----KRIGFKNMILFGLSHSMLPT 345

QY 434 -----GEV---KRRLSLYLTKIHFWLPVKMIR 458

Db 346 LLIATATLGYSGKVIDEKLYSALILTALFEAIVMSMIK 384

RESULT 4

T28867

hypothetical protein R03H4.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 03-Nov-2000

C:Accession: T28867

R:Bradshaw, H.

submitted to the EMBL Data Library, March 1996

A:Description: The sequence of C. elegans cosmid R03H4.

A:Reference number: 220534

A:Accession: T28867

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-645 <BRA>

A:Cross-references: EMBL:U50300; PIDN:AAAC48105.1; GSPDB:GN00023; CESP:R03H4.5

A:Experimental source: strain Bristol N2; clone R03H4

C:Genetics:

A:Gene: CESP:R03H4.5

A:Map position: 5

A:Introns: 35/2; 123/3; 166/3; 279/3; 320/1; 422/3; 463/1; 499/2; 527/2; 590/3

C:Superfamily: Caenorhabditis elegans hypothetical protein T26H2.7

[illegible]

Ca2+-transporting ATPase (EC 3.6.1.38) is sarcoplasmic/endoplasmic reticulum - pig N:Alternate names: Ca2+-transporting ATPase isoform 2a; calcium pump
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 22-Jun-1999
C:Accession: S04651; S15078
R:Ergermont, J.A.; Wuytack, F.; De Jaegere, S.; Nelles, L.; Casteels, R.
Biochem. J. 260, 757-761, 1989
A:Title: Evidence for two isoforms of the endoplasmic-reticulum Ca(2+) pump in pig sm
A:Reference number: S04651; MUID:89350878
A:Accession: S04651
A:Molecule type: mRNA
A:Residues: 1-997 <EGW>
A:Cross-references: EMBL:X15073; NID:g1920; PIDN:CAA33169.1; PID:g1921
A:Experimental source: stomach; smooth muscle; sarcoplasmic/endoplasmic reticulum
R:Ergermont, J.A.; Wuytack, F.; Casteels, R.
Biochim. Biophys. Acta 1088, 448-451, 1991
A:Title: Characterization of the 3' end of the pig sarcoplasmic/endoplasmic-reticulum
A:Reference number: S15077; MUID:91198156
A:Accession: S15078
A:Molecule type: DNA
A:Residues: 841-997 <EGW>
A:Cross-references: EMBL:X53754
A:Experimental source: sarcoplasmic/endoplasmic reticulum
C:Genetics:
A:Gene: SERCA2a
A:Introns: 841/1; 869/3; 914/2; 953/3; 994/1
C:Function:
A:Description: catalyzes hydrolysis of one molecule ATP coupled to translocation of t
es such as the sarcoplasmic or endoplasmic reticulum; Ca2+ pump
A:Pathway: oxidative phosphorylation

```

Qy 70 -----VLLCSKHARCFILLVFLSCGLREGRNALIAAGTGVILGHVENIFHNF 118
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 876 KEDNPDFSVCVVFESYPMTALSVLVVTIEMCNALNSL-SENSQLEMPPWENIW--- 931
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 119 KGLLDGMTNLRAKSFSTHFPLLKKYIRAIOWIYGLATPLSVFDLDLVSNQTLAVSL 175
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 932 ---LVGAIC-----LSMSLHFLIL--YVEPLPIIFQI-TPLNV-----TQMLMVLKISL 974
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
S04652
Ca2+-transporting ATPase (EC 3.6.1.38) 2, sarcoplasmic/endoplasmic reticulum
N:Alternate names: Ca2+-transporting ATPase isoform 2b; calcium pump; sarco-
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 20-Jun-2000
C:Accession: S04652; S10244; S15077
R:Egermont, J.A.; Wuytack, F.; De Jaegere, S.; Nelles, L.; Casteels, R.
Biochem. J. 260, 757-761, 1989
A:Title: Evidence for two isoforms of the endoplasmic-reticulum Ca(2+) pump
A:Reference number: S04651; MUID:89350878
A:Accession: S04652
A:Molecule type: mRNA
A:Residues: 1-1042 <EGG>
A:Cross-references: EMBL:X15074; NID:g1922; PIDN:CAA33170.1; PID:g1923
A:Experimental source: smooth muscle; sarcoplasmic/endoplasmic reticulum
R:Egermont, J.A.; Wuytack, F.; Casteels, R.

```

A:Reference number: S10244; MUID:90226293
A:Accession: S10244
A:Molecule type: mRNA

A;Residues: 989-1042 <EG>
A;Cross-references: EMBL:X16951; NID:g1924; PIDN:CAA34825.1; PID:g1364247
A;Experimental source: sarcoplasmic/endoplasmic reticulum
R;Eggenmont, J.A.; Wuytack, F.; Casteels, R.
Biochim. Biophys. Acta 1088, 448-451, 1991
A;Title: Characterization of the 3' end of the pig sarcoplasmic/endoplasmic
A;Reference number: S15077; MUID:91198156
A;Accession: S15077
A;Molecule type: DNA
A;Residues: 841-1042 <EGW>
A;Cross-references: EMBL:X53754
A;Experimental source: sarcoplasmic/endoplasmic reticulum

A:Gene: SERCA2b
A:Introns: 841/1; 869/3; 914/2; 953/3
C:Function:
A:Description: catalyzes hydrolysis of one molecule ATP coupled to translocation
es such as the sarcoplasmic or endoplasmic reticulum; Ca2+ pump
A:Pathway: oxidative phosphorylation
A:Note: membrane-bound enzyme; magnesium-dependent; P-type ATPase
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding
C:Keywords: alternative splicing; ATP; calcium transport; endoplasmic reticulum
E:60-57/Domain: calcium binding #status Predicted <CAI>
E:60-78/Domain: transmembrane #status Predicted <TM01>
F:87-107/Domain: transmembrane #status Predicted <TM02>
F:108-257/Domain: intracellular #status Predicted <INT1>
F:111-131/Domain: calcium binding #status Predicted <CA2>
F:132-238/Domain: transduction #status Predicted <TS0>
F:238-256/Domain: calcium binding #status Predicted <CA3>
F:263-279/Domain: transmembrane #status Predicted <TM03>
F:298-316/Domain: transmembrane #status Predicted <TM04>
F:308-759/Domain: intracellular #status Predicted <INT2>

F:330-505/Domain: catalytic #status predicted <PHY>
F:506-680/Domain: ATP binding #status predicted <ATP>
F:594-767/Domain: APase nucleotide-binding domain homology <ATN>
F:680-737/Domain: hinge #status predicted <HNG>

F:738-759/Domain: calcium binding #status predicted <CA5>
F:762-783/Domain: transmembrane #status predicted <TM05>
F:787-808/Domain: transmembrane #status predicted <TM06>
F:836-856/Domain: transmembrane #status predicted <TM07>

F:893-912/Domain: transmembrane #status predicted <TM08>
 F:930-949/Domain: transmembrane #status predicted <TM09>
 F:958-979/Domain: transmembrane #status predicted <TM10>
 F:1013-1029/Domain: transmembrane #status predicted <TM11>
 F:351/Active site: Asp (aspartylphosphate intermediate) #status predicted
 F:514/Binding site: ATP (lys) #status predicted

Query Match 4.1%; Score 100; DB 2; Length 1042;
 Best Local Similarity 26.0%; Pred. No. 2.5;
 Matches 46; Conservative 28; Mismatches 59; Indels 44; Gaps 12;
 QY 21 PRSP-----GWDFIOHLCVCCVALISVGLLSVAACWEL-----PSIIAARASWIITC 69
 Db 820 PRNPKEPLISGWL-FFRYIAIGCCVGAATVG---AAAWFIAADGGPRVTFFYQLSHFLQC 875
 QY 70 V-----LLCC---SKHARCFILLVFLSCGLREGNALIAAGTGVILGHVENIFHNF 118
 Db 876 KEDNPDEGVDCAVFESPYPTWALSVLVTIEMCNALNSL-SENQSLRLRPPWENIW--- 931
 QY 119 KGLLDGWTCLRAKSFISHPFLKLYTEATOWIYGLATPLSFDDLVSWNQTAVSL 175
 Db 932 ---LVGSIC-----LSNSLHFLIL--YVEPLPLFOI-TPL-----NLQWLMVLKISL 974

RESULT 10
 G86516
 dicarboxylase translocator [imported] - Chlamydothila pneumoniae (strain J138)
 C:Species: Chlamydothila pneumoniae, Chlamydia pneumoniae
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
 C:Accession: G86516
 R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
 Nucleic Acids Res. 28, 2311-2314, 2000
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
 A:Reference number: A86491; MUID:20330349
 A:Accession: G86516
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-470 <STO>
 A:Cross-references: GB:BA000008; NID:g8978580; PIDN:BAA98417.1; GSPDB:GN00142
 A:Experimental source: strain J138
 C:Genetics:
 A:Gene: ybhI
 C:Superfamily: 2-oxoglutarate/malate translocator

Query Match 4.1%; Score 99.5; DB 2; Length 470;
 Best Local Similarity 21.8%; Pred. No. 1.1;
 Matches 67; Conservative 46; Mismatches 115; Indels 79; Gaps 15;
 QY 176 FSPSHVLEQLNDSKGEVLSVLYQMATTTEVLSLQGL-----LAFAGLSVLVLGTLGFM 231
 Db 22 FSPH---PASINSNAWOLFPAIF-----TTIMGIIPQVPVPGAIAIGISTLLLTOTLT 73
 QY 232 KR-----FLGPCGKWKYNIYTRQVQDERHQRQPCVLPPLNKEERKYYIIPFTWPT 287
 Db 74 EQGLSGFHNPIAMLV-----FLSPSIAG-----IIKTGLGERIAYFFVSALGKSP 119
 QY 288 KERNLGL-----FFPLILHLCLWLVFAAVDYLLYRLIFSVMKQFOSLPFGFVHLK 339
 Db 120 -----LGLSYGLVITDFFLAPAP-----SVTARAGGLYPVWTSLSDFGS----- 161
 QY 340 LHGEKQGTQDI IHDSFNISV-FEPCNCPKPKFLLSETWPLSVILL----- 385
 Db 162 --SAEKGTDLI--GSFLIKVAQSSVITSMFLTAMAGNPLVAALAGHGVLSWVLA 217
 QY 386 -ILVMLGLSSILMLKILVSAFVSVERKRIQYHLAKLLKRSKOPF-GEVKRRRLSLY 443
 Db 218 KAALIPGLLSLFLMP---IILYKLYPPKITSCEAIRSAKLRLKMGKPLKKEETILMIF 274
 QY 444 LTKIHFV 450
 Db 275 FLVLVLW 281

RESULT 11
 G84638
 hypothetical protein At2g24610 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: G84638
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487
 A:Accession: G84638
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-690 <STO>
 A:Cross-references: GB:AE002093; NID:g4572671; PIDN:AAD23886.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g24610
 A:Map position: 2

Query Match 4.1%; Score 99.5; DB 2; Length 690;
 Best Local Similarity 24.9%; Pred. No. 1.7;
 Matches 76; Conservative 36; Mismatches 104; Indels 89; Gaps 17;
 QY 49 VAACWFLPSIIAAASWIITC---VLLCCSKHARCFILLVFLSCGL--REGNALIAAG 102
 Db 227 LGSAWILSIIERQAQTCWKAECHEKESVPLQC-----VTFDFCGTLLHRRDRNNW--QN 276
 QY 103 TGIVT--LGHVENIFHNFKGLDGMTCNLRKSFISHPFLKRYIEAIOIYGLATPLSV 160
 Db 277 TTVVFNCDPSNNIOFTFGIFADALTKNV-----VSSPFLKLYCL-W-FGL----- 322
 QY 161 FDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEVLSLQGLLAFAGL 220
 Db 323 -QNLSSYQNL-----STSTSVLETMFALIVAFGL 352
 QY 221 SL--VLLGTGLFMKFLGCGWKYNIYTRQVQDERHQRQPCVLPPLNKEERKYY 278
 Db 353 VLFALLIGN---MOTYLOSITVRLEEWRLKRR---DTEWMGHR--LIPONLRERVRRF 403
 QY 279 IIPTFWPTPKERNLGLFLP-----ILHLCIW-----LFAAVDYLLY-----RLIF 322
 Db 404 VOYKWLATRGVDEETILHSLPADLRDIQRHCLDLVRRVPLFAQMDQLDLCERLAS 463
 QY 323 SVSKQ 327
 Db 464 SLSTQ 468

RESULT 12
 A30594
 Ca2+-transporting ATPase (EC 3.6.1.38), cardiac muscle - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 16-Jun-1989 #sequence_revision 16-Jun-1989 #text_change 22-Jun-1999
 C:Accession: A30594
 R:Komuro, I.; Kurabayashi, M.; Shibasaki, Y.; Takaku, F.; Yazaki, Y.
 J. Clin. Invest. 83, 1102-1108, 1989
 A:Title: Molecular cloning and characterization of a Ca(2+)-dependent adenosin
 A:Reference number: A30343; MUID:89198060
 A:Accession: A30594
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-510 <KOM>
 A:Cross-references: GB:M25267; NID:g602485; PIDN:AAA57270.1; PID:g602486
 C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding doma
 C:Keywords: ATP; cardiac muscle; heart; hydrolase
 F:107-280/Domain: ATPase nucleotide-binding domain homology <ATN>

```
Query Match 4.0%; Score 99; DB 2; Length 510;
Best Local Similarity 25.4%; Pred. No. 1.3;
Matches 45; Conservative 28; Mismatches 60; Indels 44; Gaps 11;

QY 21 PRSP-----GWMDFIQHGLVCCVLVALISVGLLSVAACWFL-----PSIIAAAASWIITC 69
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 333 PRNPKEPLISGWL--FFRYLAIGCVGATVG---AAAWWFIAADGGPRVSYQLSHFLQC 388
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 70 -----VLLCCSKHARCFILLVFLSCGLREGRNALIAAGTGIVILGHVENIPHNF 118
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 389 KEDNPDEGVDCALFESPYMTMALSVLVTIEMCNALNSL--SENQSLLRMPWENIW--- 444
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 119 KGLLDGTCNLRAKSFISHPFLKLYEATQWIYGLATPLSVFDDLVSNQTLAVSL 175
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 445 ---LVGSIC-----LSMSLHFLIL--YVEPLPLIFQI-TPLN-----TLQWLMVLKISL 487
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
PWRBSC
Ca2+-transporting ATPase (EC 3.6.1.38), slow twitch skeletal muscle - rabbit
N;Alternate names: calcium pump, slow twitch skeletal muscle; calcium-transporting ATPase
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 18-Jun-1999
C;Accession: A01076
R;MacLennan, D.H.; Brandl, C.J.; Korczak, B.; Green, N.M.
Nature 316, 696-700, 1985
A;Title: Amino-acid sequence of a Ca(2+)+Mg(2+)-dependent ATPase from rabbit muscle sarc
A;Reference number: A01076; MUID:85296300
A;Accession: A01076
A;Molecule type: mRNA
A;Residues: 1-997 <MAC>
A;Cross-references: GB:X02814; GB:J02682; GB:M15159; NID:g1468; PIDN:CAA26583.1; PID:g14
C;Comment: This magnesium-dependent, membrane-bound enzyme catalyzes the hydrolysis of A
witch skeletal muscle sarcoplasmic reticulum; the slow twitch enzyme is low in activity.
n.
C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C;Keywords: alternative splicing; ATP; calcium transport; hydrolyase; magnesium; phosphor
F;40-57/Domain: calcium binding #status predicted <CA1>
F;60-78/Domain: transmembrane #status predicted <TM01>
F;87-107/Domain: transmembrane #status predicted <TM02>
F;107-131/Domain: transmembrane #status predicted <TM02>
F;132-238/Domain: nucleotide binding #status predicted <CA2>
F;238-256/Domain: transduction <TSD>
F;258-277/Domain: calcium binding #status predicted <CA3>
F;288-307/Domain: transmembrane #status predicted <TM03>
F;310-329/Domain: transmembrane #status predicted <TM04>
F;506-680/Domain: nucleotide binding #status predicted <CA4>
F;594-767/Domain: ATPase nucleotide-binding domain homology <NBD>
F;680-740/Domain: hinge <HNG>
F;738-759/Domain: calcium binding #status predicted <CA5>
F;760-781/Domain: transmembrane #status predicted <TM05>
F;788-808/Domain: transmembrane #status predicted <TM06>
F;831-853/Domain: transmembrane #status predicted <TM07>
F;896-915/Domain: transmembrane #status predicted <TM08>
F;927-947/Domain: transmembrane #status predicted <TM09>
F;960-981/Domain: transmembrane #status predicted <TM10>
F;351/Active site: Asp (aspartylphosphate intermediate) #status predicted
F;514/Binding site: ATP (Lys) #status predicted

Query Match 4.0%; Score 99; DB 1; Length 997;
Best Local Similarity 25.4%; Pred. No. 2.9;
Matches 45; Conservative 28; Mismatches 60; Indels 44; Gaps 11;

QY 21 PRSP-----GWMDFIQHGLVCCVLVALISVGLLSVAACWFL-----PSIIAAAASWIITC 69
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 820 PRNPKEPLISGWL--FFRYLAIGCVGATVG---AAAWWFIAADGGPRVSYQLSHFLQC 875
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 70 -----VLLCCSKHARCFILLVFLSCGLREGRNALIAAGTGIVILGHVENIPHNF 118
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 876 KEDNPDEGVDCALFESPYMTMALSVLVTIEMCNALNSL--SENQSLLRMPWENIW--- 931
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 119 KGLLDGTCNLRAKSFISHPFLKLYEATQWIYGLATPLSVFDDLVSNQTLAVSL 175
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 932 ---LVGSIC-----LSMSLHFLIL--YVEPLPLIFQI-TPLNV-----TQWLMVLKISL 974
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 14
B31981
Ca2+-transporting ATPase (EC 3.6.1.38) 2, renal - human
C;Species: Homo sapiens (man)
C;Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 22-Jun-1999
C;Accession: B31981
R;Lyttton, J.; MacLennan, D.H.
J. Biol. Chem. 263, 15024-15031, 1988
A;Title: Molecular cloning of cDNAs from human kidney coding for two alternatively sp
A;Reference number: A92673; MUID:89008384
A;Accession: B31981
A;Molecule type: mRNA
A;Residues: 1-997 <LYT>
A;Cross-references: GB:M23115; GB:J04025; NID:g184102; PIDN:AAA53194.1; PID:g306851
C;Genetics:
A;Gene: GDB:ATP2A2; ATP2B; SERCA2
A;Cross-references: GDB:119717; OMIM:108740
A;Map position: 12q23-12q24.1
C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding doma
C;Keywords: alternative splicing; ATP; hydrolyase; phosphoprotein; transmembrane prote
F;594-767/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 4.0%; Score 99; DB 2; Length 997;
Best Local Similarity 25.4%; Pred. No. 2.9;
Matches 45; Conservative 28; Mismatches 60; Indels 44; Gaps 11;

QY 21 PRSP-----GWMDFIQHGLVCCVLVALISVGLLSVAACWFL-----PSIIAAAASWIITC 69
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 820 PRNPKEPLISGWL--FFRYLAIGCVGATVG---AAAWWFIAADGGPRVSYQLSHFLQC 875
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 70 -----VLLCCSKHARCFILLVFLSCGLREGRNALIAAGTGIVILGHVENIPHNF 118
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 876 KEDNPDEGVDCALFESPYMTMALSVLVTIEMCNALNSL--SENQSLLRMPWENIW--- 931
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 119 KGLLDGTCNLRAKSFISHPFLKLYEATQWIYGLATPLSVFDDLVSNQTLAVSL 175
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 932 ---LVGSIC-----LSMSLHFLIL--YVEPLPLIFQI-TPLNV-----TQWLMVLKISL 974
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
B31982
Ca2+-transporting ATPase (EC 3.6.1.38) RS8-17 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 22-Jun-1999
C;Accession: B31982; S04269
R;Gunter-Ski-Hamblin, A.M.; Greeb, J.; Shull, G.E.
J. Biol. Chem. 263, 15032-15040, 1988
A;Title: A novel Ca(2+) pump expressed in brain, kidney, and stomach is encoded by an
ncoding Ca(2+) and other cation-transporting ATPases using an oligonucleotide probe d
A;Reference number: A92674; MUID:89008385
A;Accession: B31982
A;Molecule type: mRNA
A;Residues: 1-997 <GUN>
A;Cross-references: GB:J04023; NID:g203058; PIDN:AAA40786.1; PID:g203059
A;Experimental source: stomach
R;Lompre, A.M.; de la Bastie, D.; Boheler, K.R.; Schwartz, K.
FEBS Lett. 249, 35-41, 1989
A;Title: Characterization and expression of the rat heart sarcoplasmic reticulum Ca(2
A;Reference number: S04269; MUID:89252068
A;Accession: S04269
A;Molecule type: mRNA
A;Residues: 1-997 <LOW>
A;Cross-references: EMBL:X15635; NID:g57302; PIDN:CAA33645.1; PID:g57303
A;Experimental source: skeletal and cardiac muscle
A;Note: the authors translated the codon TGG for residue 272 as Thr, TGG for residue
C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding doma
C;Keywords: alternative splicing; ATP; calcium transport; hydrolyase; membrane protein
F;594-767/Domain: ATPase nucleotide-binding domain homology <ATN>
```

Query Match 4.0%; Score 99; DB 2; Length 997;
Best Local Similarity 25.4%; Pred. No. 2.9;
Matches 45; Conservative 28; Mismatches 60; Indels 44; Gaps 11;

Qy 21 PRSP-----GWMDFIOHLGVCCCLVALISVGLLSVAACWFL-----PSIIAAAASWIITC 69
Db 820 PRNPKEPLISGWL-FFRYLAIGCYVGAATVG---AAAWWFIAADGGPRVSFYQLSHFLOC 875
Qy 70 -----VLLCCSKHARCIFILLVFLSCGLREGRNALIAAGTGIVILGHVENIFHNF 118
Db 876 KEDNPDFEGVDCAIFESPYPMTMALSVLVTIEMCNALNSL-SENQSLLRMPWENIW--- 931
Qy 119 KLLDGMTCNLRKSFSTHFPLKKYIEAIOIYGLATPLSVFDDLVSWNQTLAVSL 175
Db 932 ---LVGSIC----LSMSLHFLIL--YVEPLPLIFOI-TPL-----NLTQWLMLVKISL 974

Search completed: January 31, 2002, 08:03:01
Job time: 142 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2002, 08:02:29 ; Search time 21.45 seconds
(without alignments)
803.378 Million cell updates/sec

Title: US-09-713-098-2
Perfect score: 2448
Sequence: 1 MGWTSQTDIFLSWEIYVS.....LPVLKMRKQMDMASADKS 470

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	102	4.2	447	1 Y341_CHLMU	O9pxk0 Chlamydia m
2	101	4.1	381	1 DHB2_RAT	Q62730 rattus norv
3	100.5	4.1	729	1 PET2_HUMAN	Q16348 homo sapien
4	100	4.1	1041	1 ATC2_CHICK	Q03669 g sarcoplas
5	100	4.1	1042	1 ATC2_PIG	P11607 s sarcoplas
6	99	4.0	997	1 ATC2_CANFA	O46674 c sarcoplas
7	99	4.0	997	1 ATC2_FELCA	Q00779 f sarcoplas
8	99	4.0	1042	1 ATC2_HUMAN	P16615 h sarcoplas
9	99	4.0	1042	1 ATC2_RABIT	P20647 o sarcoplas
10	99	4.0	1043	1 ATC2_RAT	P11507 r sarcoplas
11	99	4.0	1044	1 ATC2_MOUSE	O55143 m sarcoplas
12	96	3.9	451	1 Y069_CHLTR	O84072 chlamydia t
13	96	3.9	500	1 YJEM_ECOLI	P39282 escherichia
14	94	3.8	729	1 PET2_RABIT	P46029 oryctolagus
15	93.5	3.8	622	1 VAIL_TREPA	O83444 treponema p
16	92	3.8	858	1 CYAG_DICDI	Q03101 dictyosteli
17	91	3.7	468	1 NAH_SCHPO	P36606 schizosacch
18	91	3.7	470	1 STE3_YEAST	O67803 saccharomyc
19	91	3.7	515	1 NU2M_BETVU	P15688 beta vulgar
20	91	3.7	638	1 NUSM_PARLI	P12776 paracentrot
21	91	3.7	1381	1 ACCB_CRICR	Q09427 cricetus cr
22	90.5	3.7	424	1 ER24_SCHPO	Q09195 schizosacch
23	90.5	3.7	1418	1 CE11_CAEEL	P34641 caenorhabdi
24	90	3.7	706	1 Y006_RICPR	O9zed6 rickettsia
25	90	3.7	2327	1 CCAB_MOUSE	O55017 mus musculu
26	89.5	3.7	369	1 TAZR_HUMAN	P21731 homo sapien
27	89.5	3.7	399	1 Y507_PASMU	Q9cnc5 pasteurella
28	89.5	3.7	463	1 NARK_ECOLI	P10903 escherichia
29	89.5	3.7	471	1 NORM_CAUCR	P58163 caulobacter
30	89.5	3.7	566	1 SYRD_PSESY	P33951 pseudomonas
31	89.5	3.7	674	1 NUOL_NEIMA	Q9k1b0 neisseria m
32	89	3.6	282	1 FOCB_ECOLI	P77733 escherichia
33	89	3.6	311	1 O1D4_HUMAN	P47884 homo sapien

34	89	3.6	312	1 O1D2_HUMAN	P34982 homo sapien
35	89	3.6	374	1 Y006_BORBU	O51039 borrelia bu
36	89	3.6	497	1 NU2M_CHOCR	P48903 chondrus cr
37	89	3.6	563	1 YIM0_YEAST	P40475 saccharomyc
38	89	3.6	2336	1 CCAB_RAT	Q02294 rattus norv
39	88.5	3.6	674	1 YU41_NEIMA	Q9jx92 neisseria m
40	88.5	3.6	1203	1 NUOL_CAEEL	Q11069 caenorhabdi
41	88.5	3.6	2339	1 CCAB_HUMAN	Q00375 homo sapien
42	88	3.6	453	1 CCKR_XENLA	P70031 xenopus lae
43	87.5	3.6	576	1 CYDC_HAEIN	P45081 haemophilus
44	87.5	3.6	729	1 PET2_RAT	Q63424 rattus norv
45	87.5	3.6	1503	1 TRL2_HUMAN	O94759 homo sapien

ALIGNMENTS

RESULT 1
Y341_CHLMU
ID Y341_CHLMU STANDARD; PRT; 447 AA.
AC O9PKX0;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROBABLE METAL TRANSPORT SYSTEM MEMBRANE PROTEIN TC0341.
GN TC0341.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MoPn / Nigg;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).

CC -!- FUNCTION: PART OF AN ATP-DRIVEN TRANSPORT SYSTEM
CC TC0338/TC0339/TC0341/TC0342 FOR A METAL.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC (PROBABLE).
CC -!- SIMILARITY: BELONGS TO THE ABC-3 SUBFAMILY OF INTEGRAL MEMBRANE
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE002301; AAF73548.1; -.
CC TIGR; TC0341; -.
DR InterPro; IPR001367; HTH_DtxR.
DR InterPro; IPR001626; ABC-3.
DR Pfam; PF00950; ABC-3; 1.
DR SMART; SM00529; HTH_DTXR; 1.
DR SMART; SM00529; HTH_DTXR; 1.
KW Hypothetical protein; Transport; Transmembrane; Inner membrane;
KW Complete proteome.
FT TRANSMEM 14 34 POTENTIAL.
FT TRANSMEM 38 58 POTENTIAL.
FT TRANSMEM 69 89 POTENTIAL.
FT TRANSMEM 100 120 POTENTIAL.
FT TRANSMEM 152 172 POTENTIAL.
FT TRANSMEM 192 212 POTENTIAL.
FT TRANSMEM 233 253 POTENTIAL.
FT TRANSMEM 270 290 POTENTIAL.

cotransporter family, from human kidney.*;
 RL Biochim. Biophys. Acta 1235:461-466(1995).
 CC -!- FUNCTION: PROTON-COUPLED INTAKE OF OLIGOPEPTIDES OF 2 TO 4
 CC AMINO ACIDS WITH A PREFERENCE FOR DIPEPTIDES.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO THE PTR2 FAMILY OF TRANSPORTERS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: S78203; AAB34388.1; -.
 DR MIM: 602339; -.
 DR InterPro: IPR00109; PTR2.
 DR Pfam: PF00854; PTR2; 2.
 DR PROSITE: PS01022; PTR2_1; 1.
 DR PROSITE: PS01023; PTR2_2; 1.
 KW Peptide transport; Transport; Transmembrane; Symport; Glycoprotein.
 FT TRANSMEM 58 78
 FT TRANSMEM 88 108
 FT TRANSMEM 115 135
 FT TRANSMEM 140 160
 FT TRANSMEM 184 204
 FT TRANSMEM 218 238
 FT TRANSMEM 296 316
 FT TRANSMEM 344 364
 FT TRANSMEM 381 401
 FT TRANSMEM 588 608
 FT TRANSMEM 612 632
 FT TRANSMEM 644 664
 FT TRANSMEM 675 695
 FT CARBOHYD 7 7
 FT CARBOHYD 269 269
 FT CARBOHYD 373 373
 FT CARBOHYD 435 435
 FT CARBOHYD 472 472
 FT CARBOHYD 528 528
 FT CARBOHYD 567 567
 SQ SEQUENCE 729 AA; 81940 MW; F046073D27C063D3 CRC64;

 Query Match 4.1%; Score 100.5; DB 1; Length 729;
 Best Local Similarity 19.6%; Pred. No. 2,3;
 Matches 79; Conservative 60; Mismatches 120; Indels 145; Gaps 16;

 QY 135 SIHFPLKKYIEAOWIYGL-ATPLSVFDDLVSNQTLAVSL-----FSPSHVLEA 184
 DB 45 STAFIVNFCERFSY-YGMKAVILYFLYFLHWNEDTSTSYHAFSSLCYFPP--ILGA 101

 QY 185 QLNDS-KGEVLVLY-QWATTEVLSSLG-----QKLLAFAGLSVLVLTGLFWM 231
 DB 102 ATADSLGKFKTIISLVYLVGHVKSILGALPILGQGVVHTVLSLIGLSIALGTG--- 158

 QY 232 KRFELGPGWKYENIYITRFQVDFDERHQRPCVPLNKEERKVV-----II 280
 DB 159 --GIKFCVAAGG-----DOFEK-----HAEERTRYFVYLSINAGSLI 197

 QY 281 PTFWPTPKRKNLG-----LFFPLILHLCIWLFLA-----AVDYLL 317
 DB 198 STF-ITPMLRGDVGQGEDCYALAFGVPGLLVIALVFPAMGSKIYKPPPEGNIQAQVF 256

 QY 318 YRLIFSVMKQFOS-----LPGFEVHL 338
 DB 257 KCIWFALSNRFRKNSGDIKPRHDLWDAAEKYPQLIMDVKALTRVFLYIPLPMEWALL 316

 QY 339 KLHGEKQGTQDIHDSFNISFEENCIPKPKFLLSETWVPL----- 380
 DB 317 DOOGSRWTLQATRMNRNLGFFVLQPDQMOVNLPLVIFLPIFLDFVYIYRLVSKCGINFSS 376

QY 381 ---SVILLILVMLGLSSILMQLKILVSASFVPSVERKRIQYVLH 421
 DB 377 LRKMAVGMLACLAFAVAARVEIKINEMAPAPQPGQEVFLQVLN 420

 RESULT 4
 ATC2_CHICK STANDARD; PRT: 1041 AA.
 ID ATC2_CHICK
 AC Q03669;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE SARCOPLASMIC/ENDOPLASMIC RETICULUM CALCIUM ATPASE 2 (EC 3.6.3.8)
 DE (CALCIUM PUMP 2) (SERCA2) (SR CA(2+)-ATPASE 2) (CALCIUM-TRANSPORTING
 DE ATPASE SARCOPLASMIC RETICULUM TYPE, SLOW TWITCH SKELETAL MUSCLE
 DE ISOFORM) (ENDOPLASMIC RETICULUM CLASS 1/2 CA(2+) ATPASE).
 GN ATP2A2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS SERCA2A AND SERCA2B).
 RC STRAIN=WHITE LECHORN; TISSUE=Heart, and Brain;
 RX MEDLINE=91340754; PubMed=1831452;
 RA Campbell A.M., Kessler P.D., Sagara Y., Inesi G., Fambrough D.M.;
 RT "Nucleotide sequences of avian cardiac and brain SR/ER Ca(2+)-ATPases
 RT and functional comparisons with fast twitch Ca(2+)-ATPase. Calcium
 RT affinities and inhibitor effects".
 RL J. Biol. Chem. 266:16050-16055(1991).
 CC -!- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE HYDROLYSIS
 CC OF ATP COUPLED WITH THE TRANSLOCATION OF CALCIUM FROM THE CYTOSOL
 CC TO THE SARCOPLASMIC RETICULUM LUMEN. ISOFORM SERCA2A IS INVOLVED
 CC IN THE REGULATION OF THE CONTRACTION/RELAXATION CYCLE (BY
 CC SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + CA(2+)(IN) = ADP + PHOSPHATE +
 CC CA(2+)(OUT).
 CC -!- ENZYME REGULATION: REVERSIBLY INHIBITED BY PHOSPHOLAMBAN (PLN) AT
 CC LOW CALCIUM CONCENTRATIONS. DEPHOSPHORYLATED PLN DECREASES THE
 CC APPARENT AFFINITY OF THE ATPASE FOR CALCIUM. THIS INHIBITION IS
 CC REGULATED BY THE PHOSPHORYLATION OF PLN (BY SIMILARITY).
 CC -!- SUBUNIT: ASSOCIATED WITH PHOSPHOLAMBAN (PLN) (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. SARCOPLASMIC AND
 CC ENDOPLASMIC RETICULUM.
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS: SERCA2A/ATP2A2A AND
 CC SERCA2B/ATP2A2B (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING
 CC IN A TISSUE-SPECIFIC MANNER.
 CC -!- TISSUE SPECIFICITY: ONLY ISOFORM SERCA2A IS DETECTED IN HEART,
 CC WHILE BOTH ISOFORMS ARE EXPRESSED IN BRAIN, WITH SERCA2B BEING THE
 CC PREDOMINANT FORM.
 CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
 CC (E1-E2 ATPASES). SUBFAMILY IIA.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M66385; AAA49066.1; -.
 DR InterPro: IPR001757; E1-E2_ATPase.
 DR InterPro: IPR000695; HATPase.
 DR InterPro: IPR001454; Hydrolase.
 DR InterPro: IPR000661; Na_H_K_ATPase.
 DR Pfam: PF00122; E1-E2_ATPase; 1.
 DR Pfam: PF00702; Hydrolase; 1.
 DR PRINTS: PR00119; CATATPASE.
 DR PROSITE: PS00134; ATPASE_E1_E2; 1.
 DR KW Hydrolyase; Calcium transport; Transmembrane; Phosphorylation;
 KW ATP-binding; Metal-binding; Magnesium; Multigene family;


```

FT TRANSMEM 757 776 5 (BY SIMILARITY).
FT DOMAIN 777 786 LUMENAL (BY SIMILARITY).
FT TRANSMEM 787 807 6 (BY SIMILARITY).
FT TRANSMEM 808 827 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 828 850 7 (BY SIMILARITY).
FT TRANSMEM 851 896 LUMENAL (BY SIMILARITY).
FT TRANSMEM 897 916 8 (BY SIMILARITY).
FT TRANSMEM 917 929 9 (BY SIMILARITY).
FT TRANSMEM 930 948 10 (BY SIMILARITY).
FT TRANSMEM 949 963 10 (BY SIMILARITY).
FT TRANSMEM 964 984 10 (BY SIMILARITY).
FT TRANSMEM 985 1042 INTERACTS WITH PHOSPHOLAMBAN 1 (BY
FT TRANSMEM 370 400 SIMILARITY).
FT TRANSMEM 787 807 INTERACTS WITH PHOSPHOLAMBAN 2 (BY
FT TRANSMEM 351 351 SIMILARITY).
FT TRANSMEM 702 702 PHOSPHORYLATION (BY SIMILARITY).
FT TRANSMEM 706 706 MAGNESIUM (BY SIMILARITY).
FT TRANSMEM 304 304 2 (BY SIMILARITY).
FT TRANSMEM 305 305 2 (BY SIMILARITY).
FT TRANSMEM 307 307 2 (BY SIMILARITY).
FT TRANSMEM 309 309 2 (BY SIMILARITY).
FT TRANSMEM 767 767 1 (BY SIMILARITY).
FT TRANSMEM 770 770 1 (BY SIMILARITY).
FT TRANSMEM 795 795 2 (BY SIMILARITY).
FT TRANSMEM 798 798 1 (BY SIMILARITY).
FT TRANSMEM 799 799 1 AND 2 (BY SIMILARITY).
FT TRANSMEM 907 907 1 (BY SIMILARITY).
FT TRANSMEM 994 1042 GRECVOPATKSCFSFSACTGSGWPFVLLIMPLVWVSTDT
FT TRANSMEM 1042 AA; 114791 MW; A68EC9E41494D532 CRC64;
FT TRANSMEM 4.1%; Score 100; DB 1; Length 1042;
FT TRANSMEM 26.0%; Pred. No. 3.7;
FT TRANSMEM 46; Conservative 28; Mismatches 59; Indels 44; Gaps 12;

QY 21 PRSP-----GWMDFIQLHGVCCVALISVGLLSVAACWFL-----PSIIAAASWITC 69
Db 820 PRNPKEPLISGLW-FFRYLAIGCYVGAATVG---AAAWFIAADGGPRVTFYQLSHELQC 875

QY 70 V-----LLCC---SKHRCFILLVFLSCGLREGNRNALIAAGTGIVILGHVENIPHNF 118
Db 876 KEDNPFEGVDCAFVFPSPMTMALSVLVTIEMCNALNSL-SENQSLRPPWENIW--- 931

QY 119 KGLLDGWTCLNRAKSIHFPPLKKYIEAIOIWTYGLATPLSVFDDLVSNQTLAVSL 175
Db 932 ---LVGSIC-----LSMSLHFLIL--YVEPLIFQI-TPL-----NLQWLMLVKLSL 974

RESULT 6
ATC2_CANFA STANDARD; PRT; 997 AA.
AC 046674;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SARCOPLASMIC/ENDOPLASMIC RETICULUM CALCIUM ATPASE 2 (EC 3.6.3.8)
DE (CALCIUM PUMP 2) (SERCA2) (SR CA(2+)-ATPASE 2) (CALCIUM-TRANSPORTING
DE ATPASE SARCOPLASMIC RETICULUM TYPE, SLOW TWITCH SKELETAL MUSCLE
DE ISOFORM) (ENDOPLASMIC RETICULUM CLASS 1/2 CA(2+) ATPASE).
GN ATP2A2.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=97332675; PubMed=9188486;
RA Autry J.M., Jones L.R.;
RT *Functional Co-expression of the canine cardiac Ca2+ pump and

```

```

RT phospholamban in Spodoptera frugiperda (Sf21) cells reveals new
RL J. Biol. Chem. 272:15872-15880(1997).
CC -1- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE HYDROLYSIS
CC OF ATP COUPLED WITH THE TRANSLOCATION OF CALCIUM FROM THE CYTOSOL
CC TO THE SARCOPLASMIC RETICULUM LUMEN. ISOFORM SERCA2A IS INVOLVED
CC IN THE REGULATION OF THE CONTRACTION/RELAXATION CYCLE (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + CA(2+)(IN) -> ADP + PHOSPHATE +
CC CA(2+)(OUT).
CC -1- ENZYME REGULATION: REVERSIBLY INHIBITED BY PHOSPHOLAMBAN (PLN) AT
CC LOW CALCIUM CONCENTRATIONS. DEPHOSPHORYLATED PLN DECREASES THE
CC APPARENT AFFINITY OF THE ATPASE FOR CALCIUM. THIS INHIBITION IS
CC REGULATED BY THE PHOSPHORYLATION OF PLN.
CC -1- SUBUNIT: ASSOCIATED WITH PHOSPHOLAMBAN (PLN).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. SARCOPLASMIC AND
CC ENDOPLASMIC RETICULUM.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; SERCA2A/ATP2A2A (SHOWN HERE) AND
CC SERCA2B/ATP2A2B; ARE PRODUCED BY ALTERNATIVE SPLICING IN A TISSUE-
CC SPECIFIC MANNER (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: ISOFORM SERCA2A IS HIGHLY EXPRESSED IN HEART
CC AND SLOW TWITCH SKELETAL MUSCLE. ISOFORM SERCA2B IS WIDELY
CC EXPRESSED.
CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
CC (E1-E2 ATPASES). SUBFAMILY IIA.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U94345; AAC02263.1; -.
DR InterPro: IPR001757; E1-E2 ATPase.
DR InterPro: IPR001454; Hydrolase.
DR Pfam: PF00102; E1-E2 ATPase; 1.
DR Pfam: PF00702; Hydrolase; 1.
DR PRINTS: PR00119; CATATPASE.
DR PROSITE: PS00154; ATPASE_E1_E2; 1.
KW Hydrolase; Calcium transport; Transmembrane; Phosphorylation;
KW ATP-binding; Metal-binding; Magnesium; Calcium-binding;
KW Multigene family; Alternative splicing.
FT DOMAIN 1 48
FT TRANSMEM 49 69 CYTOPLASMIC (BY SIMILARITY).
FT DOMAIN 70 89 LUMENAL (BY SIMILARITY).
FT TRANSMEM 90 110 2 (BY SIMILARITY).
FT TRANSMEM 111 253 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 254 273 3 (BY SIMILARITY).
FT TRANSMEM 274 295 LUMENAL (BY SIMILARITY).
FT TRANSMEM 296 313 4 (BY SIMILARITY).
FT TRANSMEM 314 756 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 757 776 5 (BY SIMILARITY).
FT TRANSMEM 777 786 LUMENAL (BY SIMILARITY).
FT TRANSMEM 787 807 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 808 827 7 (BY SIMILARITY).
FT TRANSMEM 828 850 8 (BY SIMILARITY).
FT TRANSMEM 851 896 LUMENAL (BY SIMILARITY).
FT TRANSMEM 897 916 8 (BY SIMILARITY).
FT TRANSMEM 917 929 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 930 948 9 (BY SIMILARITY).
FT TRANSMEM 949 963 LUMENAL (BY SIMILARITY).
FT TRANSMEM 964 984 10 (BY SIMILARITY).
FT TRANSMEM 985 997 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 370 400 INTERACTS WITH PHOSPHOLAMBAN 1 (BY
FT TRANSMEM 787 807 SIMILARITY).
FT TRANSMEM 351 351 INTERACTS WITH PHOSPHOLAMBAN 2 (BY
FT TRANSMEM 702 702 SIMILARITY).
FT TRANSMEM 706 706 PHOSPHORYLATION (BY SIMILARITY).
FT TRANSMEM 304 304 MAGNESIUM (BY SIMILARITY).
FT TRANSMEM 305 305 MAGNESIUM (BY SIMILARITY).

```


Db 876 KDNPDFEGVDCAFSPMTALSVLTICMNCNLPSL-SENQSLRRPPHENIW--- 931
 QY 119 KGLLDGWTNCRKSPSHFPLKLYIEAIOYIGLATPLSLVDDLVSMQTLAVSL 175
 Db 932 ---LVGSIC-----LSMSLHFLIL--YVEPLPLFIQI-TPL-----NLQWLMLWLKISL 974

RESULT 8
 ATC2_HUMAN
 ID ATC2_HUMAN STANDARD; PRT: 1042 AA.
 AC P16615; P16614;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE SARCOPLASMIC/ENDOPLASMIC RETICULUM CALCIUM ATPASE 2 (EC 3.6.3.8)
 DE (CALCIUM PUMP 2) (SERCA2) (SR CA(2+)-ATPASE 2) (CALCIUM-TRANSPORTING
 DE ATPASE SARCOPLASMIC RETICULUM TYPE, SLOW TWITCH SKELETAL MUSCLE
 DE ISOFORM) (ENDOPLASMIC RETICULUM CLASS 1/2 CA(2+) ATPASE).
 GN ATP2A2 OR ATP2B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS SERCA2A AND SERCA2B).
 RC TISSUE=Kidney;
 RX MEDLINE=89008384; PubMed=2844796;
 RA Lytton J., MacLennan D.H.;
 RT "Molecular cloning of cDNAs from human kidney coding for two
 RT alternatively spliced products of the cardiac Ca2+-ATPase gene.";
 RL J. Biol. Chem. 263:15024-15031(1988).
 RN [2]
 RP VARIANTS DD.
 RX MEDLINE=99371767; PubMed=10441323;
 RA Sakuntabhai A., Burge S., Monk S., Hovnanian A.;
 RT "Spectrum of novel ATP2A2 mutations in patients with Darier's
 RT disease.";
 RL Hum. Mol. Genet. 8:1611-1619(1999).
 RN [3]
 RP VARIANTS DD, AND TISSUE SPECIFICITY.
 RX MEDLINE=99371768; PubMed=10441324;
 RA Ruiz-Perez V.L., Carter S.A., Healy E., Todd C., Rees J.L.,
 RA Steijlen P.M., Carmichael A.J., Lewis H.M., Hohl D., Itin P.,
 RA Wahlquist A., Gobello T., Mazzanti C., Reggazzini R., Nagy G.,
 RA Munro C.S., Strachan T.;
 RT "ATP2A2 mutations in Darier's disease: variant cutaneous phenotypes
 RT are associated with missense mutations, but neuropsychiatric features
 RT are independent of mutation class.";
 RL Hum. Mol. Genet. 8:1621-1630(1999).
 RN [4]
 RP VARIANTS DD THR-39; ARG-560 AND LEU-765.
 RX MEDLINE=99371769; PubMed=10441325;
 RA Jacobsen N.J.O., Lyons I., Hoogendoorn B., Burge S., Kwok P.-Y.,
 RA O'Donovan M.C., Craddock N., Owen M.J.;
 RT "ATP2A2 mutations in Darier's disease and their relationship to
 RT neuropsychiatric phenotypes.";
 RL Hum. Mol. Genet. 8:1631-1636(1999).
 RN [5]
 RP VARIANTS DD GLU-23; LYS-357; PHE-495 AND ARG-749.
 RX MEDLINE=99178263; PubMed=10080178;
 RA Sakuntabhai A., Ruiz-Perez V., Carter S., Jacobsen N., Burge S.,
 RA Monk S., Smith M., Munro C.S., O'Donovan M.C., Craddock N.,
 RA Kucherlapati R., Rees J.L., Owen M.J., Lathrop G.M., Monaco A.P.,
 RA Strachan T., Hovnanian A.;
 RT "Mutations in ATP2A2, encoding a Ca2+ pump, cause Darier disease.";
 RL Nat. Genet. 21:271-277(1999).
 CC -!- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE HYDROLYSIS
 CC OF ATP COUPLED WITH THE TRANSLOCATION OF CALCIUM FROM THE CYTOSOL
 CC TO THE SARCOPLASMIC RETICULUM LUMEN. ISOFORM SERCA2A IS INVOLVED
 CC IN THE REGULATION OF THE CONTRACTION/RELAXATION CYCLE.
 CC -!- CATALYTIC ACTIVITY: ATP + H(2O) + CA(2+)(IN) = ADP + PHOSPHATE +
 CC CA(2+)(OUT).

CC -!- ENZYME REGULATION: REVERSIBLY INHIBITED BY PHOSPHOLAMBAN (PLN) AT
 CC LOW CALCIUM CONCENTRATIONS. DEPHOSPHORYLATED PLN DECREASES THE
 CC APPARENT AFFINITY OF THE ATPASE FOR CALCIUM. THIS INHIBITION IS
 CC REGULATED BY THE PHOSPHORYLATION OF PLN (BY SIMILARITY).
 CC -!- SUBUNIT: ASSOCIATED WITH PHOSPHOLAMBAN (PLN) (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. SARCOPLASMIC AND
 CC ENDOPLASMIC RETICULUM.
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS: SERCA2A/ATP2A2/CLASS 1/HK2 AND
 CC SERCA2B/ATP2A2/CLASS 2-4/HK1 (SHOWN HERE); ARE PRODUCED BY
 CC ALTERNATIVE SPLICING. SERCA2 TRANSCRIPTS DIFFER ONLY IN THEIR
 CC 3'UTR REGION AND ARE EXPRESSED IN A TISSUE-SPECIFIC MANNER.
 CC SERCA2A IS A CARDIAC/SLOW TWITCH, MUSCLE SPECIFIC ISOFORM AND
 CC SERCA2B IS A UBIQUITOUS HOUSEKEEPING ISOFORM. SERCA2A HAS A LOWER
 CC AFFINITY FOR CALCIUM AND A HIGHER CATALYTIC TURNOVER RATE.
 CC -!- TISSUE SPECIFICITY: ISOFORM SERCA2A IS HIGHLY EXPRESSED IN HEART
 CC AND SLOW TWITCH SKELETAL MUSCLE. ISOFORM SERCA2B IS WIDELY
 CC EXPRESSED, IN SMOOTH MUSCLE AND NONMUSCLE TISSUES SUCH AS IN ADULT
 CC SKIN EPIDERMIS.
 CC -!- DISEASE: DEFECTS IN ATP2A2 ARE THE CAUSE OF DARIER'S DISEASE (DD)
 CC (ALSO KNOWN AS DARIER-WHITE DISEASE; DAR). DD IS AN AUTOSOMAL
 CC DOMINANTLY INHERITED SKIN DISORDER CHARACTERIZED BY LOSS OF
 CC ADHESION BETWEEN EPIDERMAL CELLS (ACANTHOLYSIS) AND ABNORMAL
 CC KERATINIZATION. PATIENTS WITH MILD DISEASE MAY HAVE NO MORE THAN A
 CC FEW SCATTERED KERATOTIC PAPULES OR SUBTLE NAIL CHANGES, WHEREAS
 CC THOSE WITH SEVERE DISEASE ARE HANDICAPPED BY WIDESPREAD MALODOROUS
 CC KERATOTIC PLAQUES. IN A FEW FAMILIES, NEUROPSYCHIATRIC
 CC ABNORMALITIES SUCH AS MILD MENTAL RETARDATION, SCHIZOPHRENIA,
 CC BIPOLAR DISORDER AND EPILEPSY HAVE BEEN REPORTED. STRESS, UV
 CC EXPOSURE, HEAT, SWEAT, FRICTION, AND ORAL CONTRACEPTION EXACERBATE
 CC DISEASE SYMPTOMS. PREVALENCE HAS BEEN ESTIMATED AT 1 IN 50000.
 CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
 CC (E1-E2 ATPASES). SUBFAMILY IIA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M23114; AAA53193.1; -;
 CC EMBL; M23116; AAA52757.1; -;
 CC EMBL; M23115; AAA53194.1; -;
 CC EMBL; M23278; AAA52758.1; -;
 CC EMBL; M23116; AAA52758.1; JOINED.
 CC PIR; A31981; A31981.
 CC PIR; B31981; B31981.
 CC HSC-2DPAGE; P16614; HUMAN.
 CC MIM; 108740; -;
 CC MIM; 124200; -;
 CC InterPro; IPR001757; E1-E2_ATPase.
 CC InterPro; IPR001454; Hydrolase.
 CC Pfam; PF00122; E1-E2_ATPase; 1.
 CC Pfam; PF00702; Hydrolase; 1.
 CC PRINTS; PR00119; CATATPASE.
 CC PROSITE; PS00154; ATPASE_E1_E2; 1.
 CC HydroLase; Calcium transport; Transmembrane; Phosphorylation;
 CC ATP-binding; Metal-binding; Magnesium; Calcium-binding;
 CC Multigene family; Alternative splicing; Disease mutation.
 FT DOMAIN 1 48
 FT TRANSSEM 49 69
 FT TRANSSEM 70 89
 FT TRANSSEM 90 110
 FT TRANSSEM 111 253
 FT TRANSSEM 254 273
 FT TRANSSEM 274 295
 FT TRANSSEM 296 313
 FT TRANSSEM 314 756
 FT TRANSSEM 757 776
 FT TRANSSEM 777 786
 FT TRANSSEM 787 808
 FT TRANSSEM 808 827
 FT TRANSSEM 827 847
 FT TRANSSEM 847 867
 FT TRANSSEM 867 887
 FT TRANSSEM 887 907
 FT TRANSSEM 907 927
 FT TRANSSEM 927 947
 FT TRANSSEM 947 967
 FT TRANSSEM 967 987
 FT TRANSSEM 987 1007
 FT TRANSSEM 1007 1027
 FT TRANSSEM 1027 1047
 FT TRANSSEM 1047 1067
 FT TRANSSEM 1067 1087
 FT TRANSSEM 1087 1107
 FT TRANSSEM 1107 1127
 FT TRANSSEM 1127 1147
 FT TRANSSEM 1147 1167
 FT TRANSSEM 1167 1187
 FT TRANSSEM 1187 1207
 FT TRANSSEM 1207 1227
 FT TRANSSEM 1227 1247
 FT TRANSSEM 1247 1267
 FT TRANSSEM 1267 1287
 FT TRANSSEM 1287 1307
 FT TRANSSEM 1307 1327
 FT TRANSSEM 1327 1347
 FT TRANSSEM 1347 1367
 FT TRANSSEM 1367 1387
 FT TRANSSEM 1387 1407
 FT TRANSSEM 1407 1427
 FT TRANSSEM 1427 1447
 FT TRANSSEM 1447 1467
 FT TRANSSEM 1467 1487
 FT TRANSSEM 1487 1507
 FT TRANSSEM 1507 1527
 FT TRANSSEM 1527 1547
 FT TRANSSEM 1547 1567
 FT TRANSSEM 1567 1587
 FT TRANSSEM 1587 1607
 FT TRANSSEM 1607 1627
 FT TRANSSEM 1627 1647
 FT TRANSSEM 1647 1667
 FT TRANSSEM 1667 1687
 FT TRANSSEM 1687 1707
 FT TRANSSEM 1707 1727
 FT TRANSSEM 1727 1747
 FT TRANSSEM 1747 1767
 FT TRANSSEM 1767 1787
 FT TRANSSEM 1787 1807
 FT TRANSSEM 1807 1827
 FT TRANSSEM 1827 1847
 FT TRANSSEM 1847 1867
 FT TRANSSEM 1867 1887
 FT TRANSSEM 1887 1907
 FT TRANSSEM 1907 1927
 FT TRANSSEM 1927 1947
 FT TRANSSEM 1947 1967
 FT TRANSSEM 1967 1987
 FT TRANSSEM 1987 2007
 FT TRANSSEM 2007 2027
 FT TRANSSEM 2027 2047
 FT TRANSSEM 2047 2067
 FT TRANSSEM 2067 2087
 FT TRANSSEM 2087 2107
 FT TRANSSEM 2107 2127
 FT TRANSSEM 2127 2147
 FT TRANSSEM 2147 2167
 FT TRANSSEM 2167 2187
 FT TRANSSEM 2187 2207
 FT TRANSSEM 2207 2227
 FT TRANSSEM 2227 2247
 FT TRANSSEM 2247 2267
 FT TRANSSEM 2267 2287
 FT TRANSSEM 2287 2307
 FT TRANSSEM 2307 2327
 FT TRANSSEM 2327 2347
 FT TRANSSEM 2347 2367
 FT TRANSSEM 2367 2387
 FT TRANSSEM 2387 2407
 FT TRANSSEM 2407 2427
 FT TRANSSEM 2427 2447
 FT TRANSSEM 2447 2467
 FT TRANSSEM 2467 2487
 FT TRANSSEM 2487 2507
 FT TRANSSEM 2507 2527
 FT TRANSSEM 2527 2547
 FT TRANSSEM 2547 2567
 FT TRANSSEM 2567 2587
 FT TRANSSEM 2587 2607
 FT TRANSSEM 2607 2627
 FT TRANSSEM 2627 2647
 FT TRANSSEM 2647 2667
 FT TRANSSEM 2667 2687
 FT TRANSSEM 2687 2707
 FT TRANSSEM 2707 2727
 FT TRANSSEM 2727 2747
 FT TRANSSEM 2747 2767
 FT TRANSSEM 2767 2787
 FT TRANSSEM 2787 2807
 FT TRANSSEM 2807 2827
 FT TRANSSEM 2827 2847
 FT TRANSSEM 2847 2867
 FT TRANSSEM 2867 2887
 FT TRANSSEM 2887 2907
 FT TRANSSEM 2907 2927
 FT TRANSSEM 2927 2947
 FT TRANSSEM 2947 2967
 FT TRANSSEM 2967 2987
 FT TRANSSEM 2987 3007
 FT TRANSSEM 3007 3027
 FT TRANSSEM 3027 3047
 FT TRANSSEM 3047 3067
 FT TRANSSEM 3067 3087
 FT TRANSSEM 3087 3107
 FT TRANSSEM 3107 3127
 FT TRANSSEM 3127 3147
 FT TRANSSEM 3147 3167
 FT TRANSSEM 3167 3187
 FT TRANSSEM 3187 3207
 FT TRANSSEM 3207 3227
 FT TRANSSEM 3227 3247
 FT TRANSSEM 3247 3267
 FT TRANSSEM 3267 3287
 FT TRANSSEM 3287 3307
 FT TRANSSEM 3307 3327
 FT TRANSSEM 3327 3347
 FT TRANSSEM 3347 3367
 FT TRANSSEM 3367 3387
 FT TRANSSEM 3387 3407
 FT TRANSSEM 3407 3427
 FT TRANSSEM 3427 3447
 FT TRANSSEM 3447 3467
 FT TRANSSEM 3467 3487
 FT TRANSSEM 3487 3507
 FT TRANSSEM 3507 3527
 FT TRANSSEM 3527 3547
 FT TRANSSEM 3547 3567
 FT TRANSSEM 3567 3587
 FT TRANSSEM 3587 3607
 FT TRANSSEM 3607 3627
 FT TRANSSEM 3627 3647
 FT TRANSSEM 3647 3667
 FT TRANSSEM 3667 3687
 FT TRANSSEM 3687 3707
 FT TRANSSEM 3707 3727
 FT TRANSSEM 3727 3747
 FT TRANSSEM 3747 3767
 FT TRANSSEM 3767 3787
 FT TRANSSEM 3787 3807
 FT TRANSSEM 3807 3827
 FT TRANSSEM 3827 3847
 FT TRANSSEM 3847 3867
 FT TRANSSEM 3867 3887
 FT TRANSSEM 3887 3907
 FT TRANSSEM 3907 3927
 FT TRANSSEM 3927 3947
 FT TRANSSEM 3947 3967
 FT TRANSSEM 3967 3987
 FT TRANSSEM 3987 4007
 FT TRANSSEM 4007 4027
 FT TRANSSEM 4027 4047
 FT TRANSSEM 4047 4067
 FT TRANSSEM 4067 4087
 FT TRANSSEM 4087 4107
 FT TRANSSEM 4107 4127
 FT TRANSSEM 4127 4147
 FT TRANSSEM 4147 4167
 FT TRANSSEM 4167 4187
 FT TRANSSEM 4187 4207
 FT TRANSSEM 4207 4227
 FT TRANSSEM 4227 4247
 FT TRANSSEM 4247 4267
 FT TRANSSEM 4267 4287
 FT TRANSSEM 4287 4307
 FT TRANSSEM 4307 4327
 FT TRANSSEM 4327 4347
 FT TRANSSEM 4347 4367
 FT TRANSSEM 4367 4387
 FT TRANSSEM 4387 4407
 FT TRANSSEM 4407 4427
 FT TRANSSEM 4427 4447
 FT TRANSSEM 4447 4467
 FT TRANSSEM 4467 4487
 FT TRANSSEM 4487 4507
 FT TRANSSEM 4507 4527
 FT TRANSSEM 4527 4547
 FT TRANSSEM 4547 4567
 FT TRANSSEM 4567 4587
 FT TRANSSEM 4587 4607
 FT TRANSSEM 4607 4627
 FT TRANSSEM 4627 4647
 FT TRANSSEM 4647 4667
 FT TRANSSEM 4667 4687
 FT TRANSSEM 4687 4707
 FT TRANSSEM 4707 4727
 FT TRANSSEM 4727 4747
 FT TRANSSEM 4747 4767
 FT TRANSSEM 4767 4787
 FT TRANSSEM 4787 4807
 FT TRANSSEM 4807 4827
 FT TRANSSEM 4827 4847
 FT TRANSSEM 4847 4867
 FT TRANSSEM 4867 4887
 FT TRANSSEM 4887 4907
 FT TRANSSEM 4907 4927
 FT TRANSSEM 4927 4947
 FT TRANSSEM 4947 4967
 FT TRANSSEM 4967 4987
 FT TRANSSEM 4987 5007
 FT TRANSSEM 5007 5027
 FT TRANSSEM 5027 5047
 FT TRANSSEM 5047 5067
 FT TRANSSEM 5067 5087
 FT TRANSSEM 5087 5107
 FT TRANSSEM 5107 5127
 FT TRANSSEM 5127 5147
 FT TRANSSEM 5147 5167
 FT TRANSSEM 5167 5187
 FT TRANSSEM 5187 5207
 FT TRANSSEM 5207 5227
 FT TRANSSEM 5227 5247
 FT TRANSSEM 5247 5267
 FT TRANSSEM 5267 5287
 FT TRANSSEM 5287 5307
 FT TRANSSEM 5307 5327
 FT TRANSSEM 5327 5347
 FT TRANSSEM 5347 5367
 FT TRANSSEM 5367 5387
 FT TRANSSEM 5387 5407
 FT TRANSSEM 5407 5427
 FT TRANSSEM 5427 5447
 FT TRANSSEM 5447 5467
 FT TRANSSEM 5467 5487
 FT TRANSSEM 5487 5507
 FT TRANSSEM 5507 5527
 FT TRANSSEM 5527 5547
 FT TRANSSEM 5547 5567
 FT TRANSSEM 5567 5587
 FT TRANSSEM 5587 5607
 FT TRANSSEM 5607 5627
 FT TRANSSEM 5627 5647
 FT TRANSSEM 5647 5667
 FT TRANSSEM 5667 5687
 FT TRANSSEM 5687 5707
 FT TRANSSEM 5707 5727
 FT TRANSSEM 5727 5747
 FT TRANSSEM 5747 5767
 FT TRANSSEM 5767 5787
 FT TRANSSEM 5787 5807
 FT TRANSSEM 5807 5827
 FT TRANSSEM 5827 5847
 FT TRANSSEM 5847 5867
 FT TRANSSEM 5867 5887
 FT TRANSSEM 5887 5907
 FT TRANSSEM 5907 5927
 FT TRANSSEM 5927 5947
 FT TRANSSEM 5947 5967
 FT TRANSSEM 5967 5987
 FT TRANSSEM 5987 6007
 FT TRANSSEM 6007 6027
 FT TRANSSEM 6027 6047
 FT TRANSSEM 6047 6067
 FT TRANSSEM 6067 6087
 FT TRANSSEM 6087 6107
 FT TRANSSEM 6107 6127
 FT TRANSSEM 6127 6147
 FT TRANSSEM 6147 6167
 FT TRANSSEM 6167 6187
 FT TRANSSEM 6187 6207
 FT TRANSSEM 6207 6227
 FT TRANSSEM 6227 6247
 FT TRANSSEM 6247 6267
 FT TRANSSEM 6267 6287
 FT TRANSSEM 6287 6307
 FT TRANSSEM 6307 6327
 FT TRANSSEM 6327 6347
 FT TRANSSEM 6347 6367
 FT TRANSSEM 6367 6387
 FT TRANSSEM 6387 6407
 FT TRANSSEM 6407 6427
 FT TRANSSEM 6427 6447
 FT TRANSSEM 6447 6467
 FT TRANSSEM 6467 6487
 FT TRANSSEM 6487 6507
 FT TRANSSEM 6507 6527
 FT TRANSSEM 6527 6547
 FT TRANSSEM 6547 6567
 FT TRANSSEM 6567 6587
 FT TRANSSEM 6587 6607
 FT TRANSSEM 6607 6627
 FT TRANSSEM 6627 6647
 FT TRANSSEM 6647 6667
 FT TRANSSEM 6667 6687
 FT TRANSSEM 6687 6707
 FT TRANSSEM 6707 6727
 FT TRANSSEM 6727 6747
 FT TRANSSEM 6747 6767
 FT TRANSSEM 6767 6787
 FT TRANSSEM 6787 6807
 FT TRANSSEM 6807 6827
 FT TRANSSEM 6827 6847
 FT TRANSSEM 6847 6867
 FT TRANSSEM 6867 6887
 FT TRANSSEM 6887 6907
 FT TRANSSEM 6907 6927
 FT TRANSSEM 6927 6947
 FT TRANSSEM 6947 6967
 FT TRANSSEM 6967 6987
 FT TRANSSEM 6987 7007
 FT TRANSSEM 7007 7027
 FT TRANSSEM 7027 7047
 FT TRANSSEM 7047 7067
 FT TRANSSEM 7067 7087
 FT TRANSSEM 7087 7107
 FT TRANSSEM 7107 7127
 FT TRANSSEM 7127 7147
 FT TRANSSEM 7147 7167
 FT TRANSSEM 7167 7187
 FT TRANSSEM 7187 7207
 FT TRANSSEM 7207 7227
 FT TRANSSEM 7227 7247
 FT TRANSSEM 7247 7267
 FT TRANSSEM 7267 7287
 FT TRANSSEM 7287 7307
 FT TRANSSEM 7307 7327
 FT TRANSSEM 7327 7347
 FT TRANSSEM 7347 7367
 FT TRANSSEM 7367 7387
 FT TRANSSEM 7387 7407
 FT TRANSSEM 7407 7427
 FT TRANSSEM 7427 7447
 FT TRANSSEM 7447 7467
 FT TRANSSEM 7467 7487
 FT TRANSSEM 7487 7507
 FT TRANSSEM 7507 7527
 FT TRANSSEM 7527 7547
 FT TRANSSEM 7547 7567
 FT TRANSSEM 7567 7587
 FT TRANSSEM 7587 7607
 FT TRANSSEM 7607 7627
 FT TRANSSEM 7627 7647
 FT TRANSSEM 7647 7667
 FT TRANSSEM 7667 7687
 FT TRANSSEM 7687 7707
 FT TRANSSEM 7707 7727
 FT TRANSSEM 7727 7747
 FT TRANSSEM 7747 7767
 FT TRANSSEM 7767 7787
 FT TRANSSEM 7787 7807
 FT TRANSSEM 7807 7827
 FT TRANSSEM 7827 7847
 FT TRANSSEM 7847 7867
 FT TRANSSEM 7867 7887
 FT TRANSSEM 7887 7907
 FT TRANSSEM 7907 7927
 FT TRANSSEM 7927 7947
 FT TRANSSEM 7947 7967
 FT TRANSSEM 7967 7987
 FT TRANSSEM 7987 8007
 FT TRANSSEM 8007 8027
 FT TRANSSEM 8027 8047
 FT TRANSSEM 8047 8067
 FT TRANSSEM 8067 8087
 FT TRANSSEM 8087 8107
 FT TRANSSEM 8107 8127
 FT TRANSSEM 8127 8147
 FT TRANSSEM 8147 8167
 FT TRANSSEM 8167 8187
 FT TRANSSEM 8187 8207
 FT TRANSSEM 8207 8227
 FT TRANSSEM 8227 8247
 FT TRANSSEM 8247 8267
 FT TRANSSEM 8267 8287
 FT TRANSSEM 8287 8307
 FT TRANSSEM 8307 8327
 FT TRANSSEM 8327 8347
 FT TRANSSEM 8347 8367
 FT TRANSSEM 8367 8387
 FT TRANSSEM 8387 8407
 FT TRANSSEM 8407 8427
 FT TRANSSEM 8427 8447
 FT TRANSSEM 8447 8467
 FT TRANSSEM 8467 8487
 FT TRANSSEM 8487 8507
 FT TRANSSEM 8507 8527
 FT TRANSSEM 8527 8547
 FT TRANSSEM 8547 8567
 FT TRANSSEM 8567 8587
 FT TRANSSEM 8587 8607
 FT TRANSSEM 8607 8627
 FT TRANSSEM 8627 8647
 FT TRANSSEM 8647 8667
 FT TRANSSEM 8667 8687
 FT TRANSSEM 8687 8707
 FT TRANSSEM 8707 8727
 FT TRANSSEM 8727 8747
 FT TRANSSEM 8747 8767
 FT TRANSSEM 8767 8787
 FT TRANSSEM 8787 8807
 FT TRANSSEM 8807 8827
 FT TRANSSEM 8827 8847
 FT TRANSSEM 8847 8867
 FT TRANSSEM 8867 8887
 FT TRANSSEM 8887 8907
 FT TRANSSEM 8907 8927
 FT TRANSSEM 8927 8947
 FT TRANSSEM 8947 8967
 FT TRANSSEM 8967 8987
 FT TRANSSEM 8987 9007
 FT TRANSSEM 9007 9027
 FT TRANSSEM 9027 9047
 FT TRANSSEM 9047 9067
 FT TRANSSEM 9067 9087
 FT TRANSSEM 9087 9107
 FT TRANSSEM 9107 9127
 FT TRANSSEM 9127 9147
 FT TRANSSEM 9147 9167
 FT TRANSSEM 9167 9187
 FT TRANSSEM 9187 9207
 FT TRANSSEM 9207 9227
 FT TRANSSEM 9227 9247
 FT TRANSSEM 9247 9267
 FT TRANSSEM 9267 9287
 FT TRANSSEM 9287 9307
 FT TRANSSEM 9307 9327
 FT TRANSSEM 9327 9347
 FT TRANSSEM 9347 9367
 FT TRANSSEM 9367 9387
 FT TRANSSEM 9387 9407
 FT TRANSSEM 9407 9427
 FT TRANSSEM 9427 9447
 FT TRANSSEM 9447 9467
 FT TRANSSEM 9467 9487
 FT TRANSSEM 9487 9507
 FT TRANSSEM 9507 9527
 FT TRANSSEM 9527 9547
 FT TRANSSEM 9547 9567
 FT TRANSSEM 9567 9587
 FT TRANSSEM 9587 9607
 FT TRANSSEM 9607 9627
 FT TRANSSEM 9627 9647
 FT TRANSSEM 9647 9667
 FT TRANSSEM 9667 9687
 FT TRANSSEM 9687 9707
 FT TRANSSEM 9707 9727
 FT TRANSSEM 9727 9747
 FT TRANSSEM 9747 9767
 FT TRANSSEM 9767 9787
 FT TRANSSEM 9787 9807
 FT TRANSSEM 9807 9827
 FT TRANSSEM 9827 9847
 FT TRANSSEM 9847 9867
 FT TRANSSEM 9867 9887
 FT TRANSSEM 9887 9907
 FT TRANSSEM 9907 9927
 FT TRANSSEM 9927 9947
 FT TRANSSEM 9947 9967
 FT TRANSSEM 9967 9987
 FT TRANSSEM 9987 10007

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; J04703; AAA31150.1; -;
DR EMBL; X52496; CAA36737.1; -;
DR EMBL; X02814; CAA26563.1; -;
DR PIR; S10335; PWRBMC.
DR PIR; A33881; A33881.
DR PIR; A01076; PWRBSC.
DR InterPro; IPR001757; E1-E2 ATPase.
DR InterPro; IPR000695; HATPase.
DR InterPro; IPR001454; Hydrolase.
DR InterPro; IPR000661; Na_H_K_ATPase.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR PRINTS; PR00119; CATATPASE.
DR PRINTS; PR00120; HATPASE.
DR PRINTS; PR00121; NAKATPASE.
DR PROSITE; P500154; ATPASE_E1_E2; 1.
KW Hydrolase; Calcium transport; Transmembrane; Phosphorylation;
KW ATP-binding; Metal-binding; Magnesium; Calcium-binding;
KW Multigene family; Alternative splicing.
FT DOMAIN 1 48
FT TRANSMEM 49 69
FT DOMAIN 70 89
FT TRANSMEM 90 110
FT DOMAIN 111 253
FT TRANSMEM 254 273
FT DOMAIN 274 313
FT TRANSMEM 296 333
FT DOMAIN 314 756
FT TRANSMEM 757 776
FT DOMAIN 777 786
FT TRANSMEM 787 807
FT DOMAIN 808 827
FT TRANSMEM 828 850
FT DOMAIN 851 896
FT TRANSMEM 897 916
FT DOMAIN 917 929
FT TRANSMEM 930 948
FT DOMAIN 949 963
FT TRANSMEM 964 984
FT DOMAIN 985 1042
FT TRANSMEM 1043 1042
FT DOMAIN 787 807
FT TRANSMEM 808 827
FT MOD_RES 351
FT METAL 702 702
FT METAL 706 706
FT CA_BIND 304 304
FT CA_BIND 305 305
FT CA_BIND 307 307
FT CA_BIND 309 309
FT CA_BIND 767 767
FT CA_BIND 770 770
FT CA_BIND 795 795
FT CA_BIND 798 798
FT CA_BIND 799 799
FT CA_BIND 907 907
FT VARSPLIC 994 1042
FT CONFLICT 578 578
FT SEQUENCE 1042 AA; 114704 MW; 4243836D67431575 CRC64;

Query Match 4.0%; Score 99; DB 1; Length 1042;
Best Local Similarity 25.4%; Pred. No. 4.4;
Matches 45; Conservative 28; Mismatches 60; Indels 44; Gaps 11;
21 PRSP-----GWMDFIQLHGVCLVALISVGLLSVAACWFL-----PSIIAAAASWIIC 69

Db 820 PRNKEPLISGL--FFRYLAICGYGAATVG---AAAWFIAADGPRVSFYQLSHFQC 875
Qy 70 -----VLCCSKHAKCFILLVFLSCGLRGRNALIAAGTIVILGHVENIFHNF 118
Db 876 KEDNPFDEGVDAIFESPYPMTMALSVELVTIEMCNALNSL--SENQSLLRMPWENIW--- 931
Qy 119 KGLLDGMTCNLRKASFSIHFFLLKKYIEAIOIWIYGLATPLSVFDDLVSNQTLAVSL 175
Db 932 ---LVGSIC-----LSMSLHFLIL--YVEPLPLIFOI--TPLNV-----TQWLMLVKLSL 974

RESULT 10
ATC2_RAT STANDARD; PRT; 1043 AA.
ID ATC2_RAT
AC P11507; P11508;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SARCOPLASMIC/ENDOPLASMIC RETICULUM CALCIUM ATPASE 2 (EC 3.6.3.8)
DE (CALCIUM PUMP 2) (SERCA2) (SR CA(2+)-ATPASE 2) (CALCIUM-TRANSPORTING
DE ATPASE SARCOPLASMIC RETICULUM TYPE, SLOW TWITCH SKELETAL MUSCLE
DE ISOFORM) (ENDOPLASMIC RETICULUM CLASS 1/2 CA(2+) ATPASE).
GN ATP2A2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS SERCA2A AND SERCA2B).
RC TISSUE=Brain;
RX MEDLINE=89008385; PubMed=2844797;
RA Guteski-Hamblin A.-M., Greeb J., Shull G.E.;
RT "A novel Ca2+ pump expressed in brain, kidney, and stomach is encoded
RT by an alternative transcript of the slow-twitch muscle sarcolemmal
RT reticulum Ca-ATPase gene. Identification of cDNAs encoding Ca2+ and
RT other cation-transporting ATPases using an oligonucleotide probe
RT derived from the ATP-binding site.";
RL J. Biol. Chem. 263:15032-15040(1988).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM SERCA2A).
RC TISSUE=Heart;
RX MEDLINE=89252068; PubMed=2542094;
RA Lompre A.M., de la Bastie D., Boheler K.R., Schwartz K.;
RT "Characterization and expression of the rat heart sarcolemmal
RT reticulum Ca2+-ATPase mRNA.";
RL FEBS Lett. 249:35-41(1989).
CC -1- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE HYDROLYSIS
CC OF ATP COUPLED WITH THE TRANSLOCATION OF CALCIUM FROM THE CYTOSOL
CC TO THE SARCOPLASMIC RETICULUM LUMEN. ISOFORM SERCA2A IS INVOLVED
CC IN THE REGULATION OF THE CONTRACTION/RELAXATION CYCLE (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + CA(2+)(IN) -> ADP + PHOSPHATE +
CC CA(2+)(OUT).
CC -1- ENZYME REGULATION: REVERSIBLY INHIBITED BY PHOSPHOLAMBAN (PLN) AT
CC LOW CALCIUM CONCENTRATIONS. DEPHOSPHORYLATED PLN DECREASES THE
CC APPARENT AFFINITY OF THE ATPASE FOR CALCIUM. THIS INHIBITION IS
CC REGULATED BY THE PHOSPHORYLATION OF PLN (BY SIMILARITY).
CC -1- SUBUNIT: ASSOCIATED WITH PHOSPHOLAMBAN (PLN) (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. SARCOPLASMIC AND
CC ENDOPLASMIC RETICULUM.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; SERCA2A/ATP2A2A AND
CC SERCA2B/ATP2A2B (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING
CC IN A TISSUE-SPECIFIC MANNER (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: ISOFORM SERCA2A IS HIGHLY EXPRESSED IN HEART
CC AND SLOW TWITCH SKELETAL MUSCLE. ISOFORM SERCA2B IS WIDELY
CC EXPRESSED.
CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
CC (E1-E2 ATPASES). SUBFAMILY IIA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

RESULT 12
Y069_CHLTR

QY 205 EVLSLIGQKLLA--FAG-----LSVLILGTGLFMKRFQPCG--WKYENIYITR 249
 Db 202 SGVRSVIGILLISAMFVAPPLSAROLSDRLSIFILSSIFG---GICGALCYFSVAFTC 257
 QY 250 QVQFQDERERHQPCVPLPLNKEKRRKVVILPTFTPKERKNLGLRPLILHLCL-- 306
 Db 258 QTV-----VEGKPSIILPT-----GPLVVFAGVLVFLCLIFS 291
 QY 307 -----WVLFADYLLYRLIFSVSKQ 327
 Db 292 WKTGWITRYFRKRWFLFSRDEHLLKIFWYLRQ 325

RESULT 13
 YJEM_ECOLI
 ID YJEM_ECOLI STANDARD; PRT; 500 AA.
 AC P39282;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHETICAL PROTEIN YJEM.
 GN YJEM OR B4156.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE=95334362; PubMed=7610040;
 RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
 RA Blattner F.R.;
 RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
 RT region from 92.8 through 100 minutes";
 RL Nucleic Acids Res. 23:2105-2119(1995).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
 CC (POTENTIAL).
 CC -!- SIMILARITY: BELONGS TO THE XASA FAMILY OF TRANSPORTERS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U14003; AAA97055.1; ALT_INIT.
 CC EMBL; AE000488; AAC77116.1; ALT_INIT.
 CC EcoGene; EG12475; YJem.
 CC InterPro; IPR002293; AA_rel_permease.1.
 CC InterPro; IPR002027; Amino_acid_permease.
 CC Pfam; PF00324; aa_permeases; 1.
 CC Hypothetical protein; Transmembrane; Inner membrane; Transport;
 CC Complete proteome.
 CC TRANSMEM 8 28 POTENTIAL.
 CC TRANSMEM 38 58 POTENTIAL.
 CC TRANSMEM 83 103 POTENTIAL.
 CC TRANSMEM 133 153 POTENTIAL.
 CC TRANSMEM 164 184 POTENTIAL.
 CC TRANSMEM 210 230 POTENTIAL.
 CC TRANSMEM 244 264 POTENTIAL.
 CC TRANSMEM 320 340 POTENTIAL.
 CC TRANSMEM 370 390 POTENTIAL.
 CC TRANSMEM 395 415 POTENTIAL.
 CC TRANSMEM 434 454 POTENTIAL.
 CC TRANSMEM 463 483 POTENTIAL.
 CC SEQUENCE 500 AA; 54725 MW; 9B32EBFBD20A1B4 CRC64;

Query Match 3.9%; Score 96; DB 1; Length 500;
 Best Local Similarity 22.8%; Pred. No. 3.3;

Matches 94; Conservative 48; Mismatches 123; Indels 148; Gaps 22;
 QY 7 GTDIFLSLWEIY-VSPRSPGM-----DFIQHLGVCLVALISVGLLSVAACW-FL 55
 Db 87 GTFMFSSVYIWMVSTSAKVVVPFSTFLYGSMDTQHWRIAGLEPTQVVGGLAVA--WMIL 144
 QY 56 PSIIAA-----AASWI-----ITCVLLCCSKHARCFILLVFLSGLR 92
 Db 145 VTVVASKGINKIARITAVGGIAYMCLNLVLLVITILLNGGHFAQDINFASPNPGYQ 204
 QY 93 EGRNAL-----IAAGTGIVILG-----HVENIHFNF-KGLLDGMTCLNRAKFSFIHPPLL 141
 Db 205 SGLAMLSFVFAFYAGGIEAVGGLVDKTEPNKFAKIV-----FAAIVISIGSL- 257
 QY 142 KKYIEAIOIYGLIATPLSVFDDLVSNQTL---AVSLFSPSHVLEAOLNDSKGEVLSVLY 198
 Db 258 -----AI-FLMGVST-----NMQOVLSNGSVNLGNITYVLMLKSLGTGLNALHL-- 300
 QY 199 QMATTEVLSSLGOKLLAFAGLSVLGLTGLFMKRFGLPGCGWKYENIYITRQVQFDERE 258
 Db 301 ----SPEASLSLGVWFARITGLSMFLAYTGAFETL----- 331
 QY 259 RHQORPCVPLNKEKRRKVIPT---FWPTPKERKN-LGLFELPILIH---LCIWIWFLA 311
 Db 332 -----CYSPL-----KAIQGTPKALWPEPMTRLNAMGMPSTIAMMMQCGLVTVFILLY 379
 QY 312 A-----VDVLLYRLIFSVSKQFOSLPG-----FEVHL 338
 Db 380 SFGGGTASAFFNKLTLMANVMTLPVLFALAFPPFKARQDLDRPFFVIFKTHL 432

RESULT 14
 PET2_RABIT
 ID PET2_RABIT STANDARD; PRT; 729 AA.
 AC P46029;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE OLIGOPEPTIDE TRANSPORTER, KIDNEY ISOFORM (PEPTIDE TRANSPORTER 2)
 DE (KIDNEY H+/PEPTIDE COTRANSPORTER).
 GN SLC15A2 OR PEPT2.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=96133922; PubMed=8552623;
 RA Boll M., Herget M., Wagener M., Weber W., Markovich D., Biber J.,
 RA Clauss W., Murer H., Daniel H.;
 RT "Expression cloning and functional characterization of the kidney
 RT cortex high-affinity proton-coupled peptide transporter";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:284-289(1996).
 CC -!- FUNCTION: PROTON-COUPLED INTAKE OF OLIGOPEPTIDES OF 2 TO 4
 CC AMINO ACIDS WITH A PREFERENCE FOR DIPEPTIDES.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO THE PTR2 FAMILY OF TRANSPORTERS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U32507; AAC48495.1;
 CC InterPro; IPR000109; PTR2.
 CC Pfam; PF00854; PTR2; 2.
 CC PROSITE; PS01022; PTR2.1; 1.
 CC PROSITE; PS01023; PTR2.2; 1.
 CC Peptide transport; Transport; Transmembrane; Symport; Glycoprotein.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2002, 08:01:49 ; Search time 57.59 Seconds
(without alignments)
1193.749 Million cell updates/sec

Title: US-09-713-098-2
Perfect score: 2448
Sequence: 1 MGWTSQTDIFLSWEIYVS.....LPVLKMKRKKQMDASADKS 470

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_17.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organalle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB ID	Description
1	2448	100.0	470	4 Q9H295	Q9H295 homo sapien
2	1438.5	58.8	367	11 Q9D619	Q9D619 mus musculu
3	189.5	7.7	551	4 Q9BR26	Q9BR26 homo sapien
4	168.5	6.9	498	11 Q9D611	Q9D611 mus musculu
5	158.5	6.5	684	5 Q9VU52	Q9VU52 drosophila
6	122.5	5.0	689	2 Q9KY46	Q9KY46 streptomyc
7	113.5	4.6	351	5 Q17520	Q17520 caenorhabdi
8	104	4.2	320	2 Q9RNE1	Q9RNE1 bacillus an
9	104	4.2	401	10 Q9FLR8	Q9FLR8 arabidopsis
10	104	4.2	1971	4 Q9NTT5	Q9NTT5 homo sapien
11	103	4.2	638	5 Q917H4	Q917H4 drosophila
12	102.5	4.2	407	2 Q9ZCV6	Q9ZCV6 rickettsia
13	102	4.2	391	2 Q9PLZ4	Q9PLZ4 campylobact
14	102	4.2	536	2 Q06762	Q06762 mycoplasma
15	101	4.1	635	2 Q9F443	Q9F443 streptococc
16	101	4.1	645	5 Q21688	Q21688 caenorhabdi
17	99.5	4.1	470	2 Q9JS55	Q9JS55 chlamydia p
18	99.5	4.1	690	10 Q9SJA4	Q9SJA4 arabidopsis
19	99	4.0	510	11 Q63080	Q63080 rattus norv

20	98.5	4.0	578	5 Q9VCG4	Q9VCG4 drosophila
21	98	4.0	363	5 Q9N5D1	Q9N5D1 caenorhabdi
22	98	4.0	598	2 Q50866	Q50866 myxococcus
23	97.5	4.0	1065	2 Q31501	Q31501 bacillus su
24	97	4.0	614	2 Q31615	Q31615 bacillus su
25	97	4.0	626	2 Q51574	Q51574 borrelia bu
26	96.5	3.9	732	5 Q18050	Q18050 caenorhabdi
27	96	3.9	405	2 Q24865	Q24865 helicobacte
28	96	3.9	700	5 Q9UAB9	Q9UAB9 leishmania
29	96	3.9	1465	3 Q9P5N0	Q9P5N0 schizosacch
30	95.5	3.9	658	5 Q9U3N1	Q9U3N1 caenorhabdi
31	95	3.9	486	2 Q9ZDP4	Q9ZDP4 rickettsia
32	94	3.8	298	8 Q9B8G3	Q9B8G3 heterodoxu
33	94	3.8	899	11 Q9Z2P2	Q9Z2P2 mus musculu
34	94	3.8	1304	10 Q9SUK4	Q9SUK4 arabidopsis
35	93.5	3.8	485	11 Q9ESQ5	Q9ESQ5 rattus norv
36	93.5	3.8	533	13 Q9Y125	Q9Y125 ambystoma t
37	93.5	3.8	725	5 Q16382	Q16382 caenorhabdi
38	93.5	3.8	921	2 Q9ZD78	Q9ZD78 rickettsia
39	93	3.8	298	1 Q9Y158	Q9Y158 pyrococcus
40	93	3.8	337	4 Q14804	Q14804 homo sapien
41	93	3.8	408	2 Q9HX29	Q9HX29 pseudomonas
42	93	3.8	437	2 Q99Z61	Q99Z61 streptococc
43	93	3.8	484	10 Q9SVJ0	Q9SVJ0 arabidopsis
44	93	3.8	933	2 Q9I4S0	Q9I4S0 pseudomonas
45	93	3.8	1451	4 Q9BQ06	Q9BQ06 homo sapien

ALIGNMENTS

RESULT 1

Q9H295 ID Q9H295 PRELIMINARY; PRT; 470 AA.
AC Q9H295;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE DC-SPECIFIC TRANSMEMBRANE PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hartgers F.C., Vissers J.L.M., Looman M.W.G., van Zoelen C.,
RA Huffin C., Figdor C.G., Adema G.J.;
RT "PC-STAMP, a novel multimeric membrane-spanning molecule preferentially
RT expressed by dendritic cells."
RL Eur. J. Immunol. 0:0-0(2001).
DR EMBL: AF305068; AAG39167.1; -;
DR InterPro: IPR001211; PLP_A2.
DR PROSITE: PS00118; PA2_HIS; UNKNOWN_1.
KW Transmembrane.
SQ SEQUENCE 470 AA; 53392 MW; EA2B858FD2C7560C CRC64;

Query Match 100.0%; Score 2448; DB 4; Length 470;
Best Local Similarity 100.0%; Pred. No. 3.6e-190;
Matches 470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGWTSQTDIFLSWEIYVS PRSGWMDFIQHLGVCCLVALLSVGLLSVAACWFLPSIIA 60
|||||
Db 1 MGWTSQTDIFLSWEIYVS PRSGWMDFIQHLGVCCLVALLSVGLLSVAACWFLPSIIA 60
|||||
QY 61 AAASWITCVLLCCSKHARCIFLLVFLSCGLREGNRLIAAGTGVILGHVENIFHNFKG 120
|||||
Db 61 AAASWITCVLLCCSKHARCIFLLVFLSCGLREGNRLIAAGTGVILGHVENIFHNFKG 120
|||||
QY 121 LLDGWTNCRAKSFIHPLKLYIEATQWIYGLATPLTSVFDLVSQNTLAVSLFSPSH 180
|||||
Db 121 LLDGWTNCRAKSFIHPLKLYIEATQWIYGLATPLTSVFDLVSQNTLAVSLFSPSH 180
|||||

```
QY 181 VLEAOLNDSKGEVLSVLYOMATTEVLSLGGOKLLAFAGLSLVLLGTGLPKRFLGPCGW 240
Db 181 VLEAOLNDSKGEVLSVLYOMATTEVLSLGGOKLLAFAGLSLVLLGTGLPKRFLGPCGW 240
QY 241 KYENIYITROFQVQDERERHQORPCVLPNKEERRKVIIPTWPTPKERKNLGLFLPI 300
Db 241 KYENIYITROFQVQDERERHQORPCVLPNKEERRKVIIPTWPTPKERKNLGLFLPI 300
QY 301 LIHLCIWLFAAVDYLLRLIFSVSKQFSLPGFEVHLKLHGKQGTQDIIHDSSEFNISV 360
Db 301 LIHLCIWLFAAVDYLLRLIFSVSKQFSLPGFEVHLKLHGKQGTQDIIHDSSEFNISV 360
QY 361 FEPNCIPKPFILSETWVPLSVLLILVMLGLSSILMQLKILVSASFYSVERKRIQYL 420
Db 361 FEPNCIPKPFILSETWVPLSVLLILVMLGLSSILMQLKILVSASFYSVERKRIQYL 420
QY 421 HAKLLKRSKQPLGEVKRRRLSLYTKIHFWLPVLMIRKKOMDASADKS 470
Db 421 HAKLLKRSKQPLGEVKRRRLSLYTKIHFWLPVLMIRKKOMDASADKS 470

RESULT 2
Q9D619 ID Q9D619 PRELIMINARY; PRT; 367 AA.
AC Q9D619;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 4833414107RIK.
GN 4833414107RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=HEAD;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi S., Fukuda S.,
RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK014697; BAB29508.1; -
DR MGD; MGI:1923016; 4833414107RIK.
SQ SEQUENCE 367 AA; 42104 MW; 964ADCF85B543B1 CRC64;
```

```
Query Match 58.8%; Score 1438.5; DB 11; Length 367;
Best Local Similarity 58.8%; Pred. No. 1.4e-108;
Matches 273; Conservative 39; Mismatches 53; Indels 103; Gaps 1;

QY 1 MGWTSGDTIFSLWEIYVSPRPGWMDFTQHLGVCCLVALISVGLLSVAACWFLPSTIA 60
Db 1 MRWTGTSIFLWGTIVYVPRSPWLDFTQHLGVCCLVALISVGLLSVAACWFLPSTIA 60
QY 61 AAASWIITCVLLCCSKHARCFILLVFLSCGLREGNALLAAGTGIVILGHVENIFNPKG 120
:: :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 61 LSSVWMITCVFLCCSKRARCIFILLVFLSCGLREGNALLAAGTGIVILGHVENIFNFRG 120
QY 121 LDGTCNLRAKSFISHFPLKKYIEAIQWYIGLATPLSVFDDLVSNQTLAYSLFSPSH 180
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 LLDSTCNLRAKSFVHFPLKKRYTEAIQWYIGLATPLNLFDDLVSNQTLVVSFLSPSH 180
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 VLEAOLNDSKGEVLSVLYOMATTEVLSLGGOKLLAFAGLSLVLLGTGLPKRFLGPCGW 240
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 ALEAHMNDTRGVLGVLLHHMVYTTLLTSVGOKLLALAGLLLVSTGLFKRFLGPCGW 240
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 KYENIYITROFQVQDERERHQORPCVLPNKEERRKVIIPTWPTPKERKNLGLFLPI 300
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 KYENIYITROFQVQDERERHQORPCVLPNKEERRKVIIVPSQLTPKERTKTLGLFLPV 300
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 LIHLCIWLFAAVDYLLRLIFSVSKQFSLPGFEVHLKLHGKQGTQDIIHDSSEFNISV 360
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 LTYLYMWVLFVFAAVDYLLRLISSMKNQFQSLPGLEVHLKLHGE----- 343
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 FEPNCIPKPFILSETWVPLSVLLILVMLGLSSILMQLKILVSASFYSVERKRIQYL 420
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 ----- 343
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 HAKLLKRSKQPLGEVKRRRLSLYTKIHFWLPVLMIRKKOMDASAD 468
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 344 ----- 365
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
Q9BR26 ID Q9BR26 PRELIMINARY; PRT; 551 AA.
AC Q9BR26;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE DJ257E24.3 (NOVEL PROTEIN) (FRAGMENT).
GN DJ257E24.3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd D.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL034424; CAC36102.1; -
FT NON_TER 1
SQ SEQUENCE 551 AA; 60104 MW; 38A47A2C6B9546AE CRC64;
```

```
Query Match 7.7%; Score 189.5; DB 4; Length 551;
Best Local Similarity 20.6%; Pred. No. 3.1e-07;
Matches 102; Conservative 71; Mismatches 201; Indels 121; Gaps 19;

QY 12 LSLWEIYVSPRPGWMDFTQHLGVCC---LVALISVGLLSVAAC-----WFL----- 55
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6 LGFWKA-LAPLOAAMDFAFSQVPASCGQLLTQLLCCASLAAAAGLYVHWSLILLYPPG 64
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 56 PSIIIAASWIITCVLLCCSKHARCFILLVFLSCGLREGNALLAAGTGIVILGHVENIF 115
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65 PSAMVATVCGLLVFLSVLPVPRVCLFALSVPITLGMGEGRRLISYSTATLATAVNPVL 124
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 116 HNFKGLLDGTCNLRAKSFISHFPLKKYIEAIQWYIGLATPLSVFDDLVSNQTLAVSL 175
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 125 ANVGAAGQVLRVCTEGS-----LESLLNTHQHAAASRALGPTGQAGSRGLT----- 171
QY 176 FSPSHVLEAQLNDSKGEVLSVLYOMATTEVLSLGG--QKLLAFAGLSLVLLGTGLPKMR 233
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 172 -----FEADNGS-----AFYHLMLRVTTQVLEDFSCLESLSARAALGTORVVVTGLF 221
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 234 FLGPCGW-----KYENIYITROFVQ--FDERERHQORP-----CVLPNKEER 274
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 222 LLVESAWYLRCVLTDLRFDNIVATQTLQRLAQQAQATHLLAPPTTLLQAAQLRLSQEEL 281
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

QY 275 RYVLIPTFWPTPKERKNGLGFPLILHLCIWLFAAVDYLLYRLIFSVSQFQSLPGF 334
Db 282 LSCLL-----RGLLAL-LLVATAVAV---ATDHVAFLLAQAIVDWAQKLPTV 325
QY 335 EYHLKHGEKOGTDIIHDSSEFNISVFEP-----NCIPKPKFL-----LSETWVP 379
Db 326 PI-----TLTVKYDVAVTVLGFIPFLFNQLAPESPFLSVHSSYQWELRLTSARCP 375
QY 380 L-----SVILLIUMGLSSILMQLKILVSASFYSVEKRKROYLHAKLL 425
Db 376 LLPARRPRAAAPLAAGALLAGLQALLAGTAVLFAVARRLRHATAAASFFTAQEARVRVHLHARLQ 435
QY 426 KRKSQ-----PLGE 435
Db 436 KRHRHQOQLPLGD 450
RESULT 4
Q9D611 PRELIMINARY; PRT; 498 AA.
ID Q9D611
AC Q9D611
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE 4833422F24RIK PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=HEAD;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kado K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner N., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., King B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sakai H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK014751; BAB29531.1; -.
DR MGD; MGI:1921864; 4833422F24RIK.
SQ SEQUENCE 498 AA: 54573 MW; FD27C03258923B55 CRC64;

Query Match 6.9%; Score 168.5; DB 11; Length 498;
Best Local Similarity 20.4%; Pred. NO. 1.4e-05;
Matches 103; Conservative 71; Mismatches 186; Indels 145; Gaps 20;

QY 15 WEIY-----VSPRSGWMDFIQHLGVCCVALLISVGLLSVAACWFLPSITIAAAA-SWI 66
Db 16 MFWRLGICKAVVPLQAAWAKFQBPVAPSCNELLTQLLL-----CVSLASLAGLHHL 70
QY 67 ITCVL-----LC-----CSKHARCFTLLVFLSCGLREGRNALIAAGTGV 106
Db 71 VSLQLYPLGPPALVTSGLGLVSLGLVPLPPIRCFLVLSVPLTGLSKQGRRLLSYAANL 130
QY 107 ILGHRVENIFHNPKGLDGMTCNLRKFSIHFPPLKKYIEATQWYIGLATPLSFDDLYS 166

Db 131 AVAVPVLGNVRAAGQVLSCTEGS-----LESLLNTTYQLRQAARELGPASRAGS 182
QY 167 WNOTLAVLSFSPSHVLEAQLNDSKEVLSVLYQMATTTEVLSSL-GQKLLAPAGLSVL 225
Db 183 RSLTFEY-----EGKGSFR-LHMHITITOEILDEFSGLEFLAARALGTQRV 227
QY 226 GTGLEMKRFLGPCW-----KYENIXITROFQVDERERHQ-----QRPCVL 267
Db 228 VTGLFLGLLGESAWYLHRYLTDLRFNINAYATRLV-----ROLAQAAGATHLLTSPPPWLL 283
QY 268 -----PLNKERRKYVIPTFWPTPKERKNGLGFPLILHLCIWLFAAVDYLLYRLIF 322
Db 284 QTAQPKLSRELLSCLL-----RGLLAL-LLVATAVTV---ASDYGAFLLAQ 327
QY 323 SVSKQFQSLGFEVHLKHGEKOGTDIIHDSSEFNISVFEPNCIPKP----- 369
Db 328 AAVAAQKLTVP-----TLTVKYDASYKVLDFILFVLNQPVPVESFASMQRSF 377
QY 370 -----KFLSETWVP-----LSVILLIUMGLSSILMQLKILVSASFYSVE 413
Db 378 QWELRFTPHDCHLPQAQPPRVTAALAAAGALLAGLQALLAGTAVLQAYAWRLRHTTAAASFPDQE 437
QY 414 RKROYLHAKLLKKR-----SKOP 432
Db 438 ARRLSHLQARLQRRHNSDHLNKP 462
RESULT 5
Q9VU52 PRELIMINARY; PRT; 684 AA.
ID Q9VU52
AC Q9VU52
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CG11281 PROTEIN.
GN CG11281.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu A., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler K., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jallil M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kamali B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

```
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT The genome sequence of Drosophila melanogaster.6;
RL Science 287:2185-2195(2000).
DR EMBL: AE003538; AAF49838.1; -.
DR FlyBase: FBgn0036347; CG11281.
DR InterPro: IPR001841; znf_ring.
DR SMART: SM00184; RING.1.
SQ SEQUENCE 684 AA; 79599 MW; 504526517354147C CRC64;

Query Match 6.5%; Score 158.5; DB 5; Length 684;
Best Local Similarity 17.1%; Pred. No. 0.00013;
Matches 91; Conservative 83; Mismatches 157; Indels 201; Gaps 17;

QY 66 IITVCLCCSKHARCFTLLVFLSCGLREGNRLIAAGTGIVILGHVNIHFNFKGLLDGM 125
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 126 TC-----NURAKSFHPLPKKYIEAIOI----- 151
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 134 VCTVLTYNLSKTRFDLMAKPFNTLKHMRGDVEIRHTFELQELVLDLKYAVENSIE 193
QY 152 ---YG-----LATPLSV----- 160
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 194 DEKYGDNKTPIYRWGRETSMNVSEIGNGKELPTPAAVQERFORNMRNRCKHQLRS 253
QY 161 -----FDDLVS---WNQTLAV-----SLFSPSHVL----- 182
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 254 HRACLEVRNGYRKCTNFNFSAIKALCWYRVDIICELDFGNPKICDPSAVVPNGF 313
QY 183 -----EAOLNDSKGEVLSLYOMATTEVLSLG----- 211
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 314 ETYVELLKAEKLFDNSSQIV-VNVEIKDQFAKSQLSAERTGOAFKEDFERQKRIFN 372
QY 212 -----OKLLAFAGLSVLTLGTLMKRFGLPCGWKYENIYITROFQDERERHQPCV 266
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 373 VMGILQKILCLFMRWVYVSYIVK-YLNDV--EFDNFYITKFKVHDQRRKEQRI 429
QY 267 LPLNKEERKYLIPTFWPTPKRNKGLFELPILHLCLTWLFAAAYDLYLLRILFVS 326
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 430 LPLRTYKSKYIDVD-----H-----IFSRTTHHESTTIVRKRK 463
QY 327 -QFQSLPFGFEVHLKHGEKQGT---QDIHDSFN-----SVFEPNCIPKPKFL 376
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 464 IVYQDGEHEVRFNISVGMARLLRTMH--NFNIHEKVSTSLSNKECLPNAHVLP 521
QY 377 WVPVLSVILLVLMGLSSILMOLKILVSASFYSPVSKERQIYLHAKLLKKR 428
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 522 YYQILILLIILVLIYQSTFLMRMRVICSFYFKKREKQRIILFLYILNRNR 573

RESULT 6
QYK146 PRELIMINARY; PRT; 689 AA.
AC Q9KY46;
AC Q9KY46;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PUTATIVE INTEGRAL MEMBRANE TRANSPORT PROTEIN.
GN SCC8A.02c.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN-A3(2);
RA Brown S.P., Harris D.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RL the 8 Mb Streptomyces coelicolor A3(2) chromosome."7;
RM Mol. Microbiol. 21:77-96(1996).
DR EMBL: AL356892; CAB92820.1; -.
DR InterPro: IPR001958; TCR_tetA.
DR PRINTS: PR01035; TCR_tetA.
DR PRINTS: PR01036; TCR_tetA.
SQ SEQUENCE 689 AA; 73450 MW; 4DA0FDB0BCA10EBDA CRC64;

Query Match 5.0%; Score 122.5; DB 2; Length 689;
Best Local Similarity 21.7%; Pred. No. 0.11;
Matches 80; Conservative 48; Mismatches 108; Indels 133; Gaps 20;

QY 21 PRSPGMDPIQHLGVCCCLVALISVGLLSVA-----ACWFLPSIIAAA-----SWITCV 70
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 18 PAAPGEREQVSS-GVLVSGALLGLMLAALDQTIIVSTALPTIVSDGLGLEHLSWVYAT 76
QY 71 LLCCSK-----HARCFILL--VEL---SCGLREGNRLIA-----AGTC 104
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 77 LLAATAATPLMKLGQYGRKKLFQALIGFLVGSALCGIAGMGQLIAFRALQGLGGG 136
QY 105 IVIL-----GHVENIFHNFKGLDGMTCNLRAKSFHSFPLPKKYIEAIO 150
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 137 LMVLSMAIVGDLVPPRGRYQGLF---GAVFGAT-----SVLGPLG----- 176
QY 151 IYGLATPLSVFDDVSNQOTLAVSLFSPSHVLEAOLNDSKGEVLSVLY-OMATTEVLSS 209
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 177 -----GVTEHLSWRWVFYINL--PIGVVALAV-----IAAVLHIPRRTTRHVIDY 220
QY 210 LGOKLLAFAGLSVLTLGTLMKRFGLPCGWKYEN-----IYITROFQVQDERERHQ 262
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 221 LGTLIASVATSLVLVAS-----LGTTWAWSSPQIIGLAVLAVVLVLFVAVERRAA 273
QY 263 RCVLPLNKEERKYVI-----IPTF-----WPTPKRNKGLGFF 297
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 274 EP-VLPLKLFVRVTFALSNAVISFVGFAMFGAMTYLPTFLQVVRGVTPMTS-----GVYM 327
QY 298 LPILHLICI 306
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 328 LPMVFGLL 336

RESULT 7
QYK1520 PRELIMINARY; PRT; 351 AA.
AC O17520;
AC O17520;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ZC132.7 PROTEIN.
GN ZC132.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
```



```
SQ SEQUENCE 401 AA; 45802 MW; F8EEB5620F92E492 CRC64;

Query Match
Best Local Similarity 4.2%; Score 104; DB 10; Length 401;
Matches 77; Conservative 56; Mismatches 111; Indels 86; Gaps 19;

QY 144 YIEAIOIYGLATPLSVF--DDLVSNNOTLAVLSFSPSHVLEAQLNDSKEVLSVLYQMAT 202
Db 53 YAAQAI-WIHGFACVLELMAEPYILSQTLL-----LKLQLR-LIVETVATLARCVT 100
QY 203 --TTEVLSSLGOKLLAFAGLSLVLLGTGLFMK--RFLGPGC-WKYENIYIIR--QFVQFD 255
Db 101 LYSLLVMOTNNKVIIFA-USQAYAGSLFTIGYWAYFLICGVYRSSNLFPRPGNFMDFD 159
QY 256 E-----RERHQ-----RCPVLPNKEERKYYIIPFETWPTPKER 290
Db 160 NDLKMKCLFTFQSFERKLLIQEVLKTPGSLVVRVFLPF---ESSYIIFARFASGDQOE 216
QY 291 KN--LGLFF-----LPLIHLCIWLFAADVLYLRLIF-----SVSKQFQSLPGF 334
Db 217 RNKKGIVLTVALKLVILIGLIFMAFGPSYSLSIRLLYGEKWSDEASLALQFYCL--Y 274
QY 335 EVHKLHGEKO-----GTQDIHDSFNSISVFPNCIPKPKFLLSETWVPLSVILLILV 388
Db 275 IIVLAMNGTSEAFUHVAGTKNELERSNDMLIF-----SLIYVALNILLIRSA 322
QY 389 -MLGSLLSILMOLKILVSAFSPVVERKRI 417
Db 323 GAIGLIWANSLSWQILLSLITLISEKTI 352

RESULT 10
Q9NTT5 ID Q9NTT5 PRELIMINARY; PRT; 1971 AA.
AC Q9NTT5;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE DJ202D23.2 (NOVEL PROTEIN) (FRAGMENT).
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Sehra H.;
RL EMBL; AL121716; CAB86625.1; -.
FT NON_TER 1
SQ SEQUENCE 1971 AA; 221833 MW; 21B0E98EB8B37FB8 CRC64;

Query Match
Best Local Similarity 4.2%; Score 104; DB 4; Length 1971;
Matches 107; Conservative 77; Mismatches 160; Indels 184; Gaps 28;

QY 6 SGTDFILSLWEIYSPSPGMDFTQHL---GVCC-----LVALISVGLLSVAACWFLPSI 58
Db 1188 TGFDFVSDLE-HISPHQP--MTSLQYLHAQPIITCQGMFLCAVIR-ALHQHCACRKHPOW 1243
QY 59 IAAASWTITCVLLCCSHKARCFILLVFLSCGLREGRNALIAAGTIVLGHVENIFHNF 118
Db 1244 IG-----LITSTLPMGK-----VQIRVVVSVTLQLCRN-----LDNLIOQY 1280
QY 119 K-----GLDGMTCNLRAKSFSIHFPLLLKKYIEAIOIYGLA 155
Db 1281 KYETGLSDSRPLWASIIPPDMILTLLEGITA-----IIHYCLL----- 1319
QY 156 TP4SVFDDLVSNNOTLAVLSFSPSHVLEAQLNDSKEVLSVLYQMATTEVLSSL----- 210
Db 1320 -----DPTQYHQ-LLVSV-DQKHLFEA-----RSGILSLHMTMSSVTLWLSILHQAD 1366
```

```
QY 211 -GOKLLAFAGLSLVLLGTGLF-----MKRFLGPGCGWKYENIYITROVFOFDERERHQ-- 261
Db 1367 SSEKMTIAASASLTITNLGATKNLQOILLELLGPISNHNHGVHFAALAFVWNNERRQKTT 1426
QY 262 QRPCVLPNKEER-----RKVYIITFWPT-----PKERNLGLFFLPIL 301
Db 1427 TRTKVIPAAASEQQLLVELVRSISVMRAETVIQTVKEVLKOPPAIAIKDKHLSL----- 1480
QY 302 IHLCIWLFAADVLYLRLIFSVSKOF-QSLPGFEVHLKLGHEKQGTQDIHDSFNSISV 360
Db 1481 -EVCMLQFF-----YAIQVRVLEFIMKNPS-----LENKKQORDQDQTHKIVDAIGA 1528
QY 361 FEPNCI-----PKPKFLLSETWVPLSVILLILVMLGSLLSILMOLKILVSA 407
Db 1529 IAGSSLEQTTWLRNRLNEVKPSPKIMVDGTNLESDVEGILVKHLGLL-SVLAHLDMV-- 1584
QY 408 FVPSVERKRIQYLHAKLLKRSKQPLGEVRRRLSLYTKTHFWLPUVK 455
Db 1585 FY-SDEKERV-----IPLLNMIMHYVVPYLR 1609

RESULT 11
Q9I7H4 ID Q9I7H4 PRELIMINARY; PRT; 638 AA.
AC Q9I7H4;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
CG9717 PROTEIN.
GN Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
```

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL; AF003774; AAG22176.1; -.
 DR FlyBase; FBgn0039789; CG9717.
 DR InterPro; IPR002645; STAS.
 DR InterPro; IPR001902; Sulfate_transp.
 DR Pfam; PF01740; STAS; 1.
 DR Pfam; PF00916; Sulfate_transp; 1.
 SQ SEQUENCE 638 AA; 69359 MW; C538F3BE5CB84F CRC64;

Query Match 4.2%; Score 103; DB 5; Length 638;
 Best Local Similarity 22.3%; Pred. No. 3.8;
 Matches 80; Conservative 54; Mismatches 113; Indels 112; Gaps 20;
 QY 1 MGIWSTGTDIFLSLWEIYVSP-RSPQWMDFIQHLGVCCLVALISVGLLSVAACWFLPSII 59
 Db 213 LGIKTSG-WTFQWLSIVNDIHNISWPDFI--LGIVCITLLLS--LRALASCTLGPKEG 267
 QY 60 AAAAIIITCVLLCCSKHARCFILVFLSCGLREGRNALIAAGTGIVLGHVENIHFNFK 119
 Db 268 KITAQKLLFG-----IPWTIG--TARNALLVCGT--AGLCYW--LFVNGK 306
 QY 120 GLLDGWTCLNRAKSFHPLKLYIEATOWI-YGIATPLSVFDDLVSNQTLAVLSLSP 178
 Db 307 -----ENLVKTVGFVKGLP-----SFQPP 326
 QY 179 SHVLEAQLNDSKEVLSVLYQMATTEVLSLQKLLAFAGLSVLVLTGCLFNMKRL--- 235
 Db 327 PFHMDAVNNTTGEVLQ---EAQSFQWDMVSTLGSGLIVVP--LIALLETMAVVOAADGK 381
 QY 236 -----GPCGKMYENITITQVQFDERERHQPCVL-----PLNKEERRY 277
 Db 382 PTDATQELLTASGVN-----VANSFVQGLRNSGGIARGAILNASCVRTQLNLTYSVI 434
 QY 278 VIITFPWTPPKRNKLGFLFLILHLCIWLFAAAYDILL-YRLFSV--SKOFQSLPG 333
 Db 435 VIATLLYLP-----CFYIIPKAALASI--IAAVIFVQYRVIKPMWHSKKTDLIPG 485

RESULT 12
 Q9ZCV6 PRELIMINARY; PRT; 407 AA.
 AC Q9ZCV6;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DE BICYCLOWYCIN RESISTANCE PROTEIN (BCR1).
 GN RP603.
 OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=782;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MADRID E;
 RX MEDLINE=99039499; PubMed=9823893;
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
 RA Sacheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.:
 RT "The genome sequence of *Rickettsia prowazekii* and the origin of
 mitochondria".;
 RL Nature 396:133-140(1998).
 DR EMBL; AJ235272; CAAL15047.1; -.
 DR InterPro; IPR003662; sub_transporter.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 407 AA; 45749 MW; 8A5EC5A0BEEF6EC6 CRC64;

Query Match 4.2%; Score 102.5; DB 2; Length 407;
 Best Local Similarity 16.3%; Pred. No. 2.4;
 Matches 79; Conservative 79; Mismatches 131; Indels 195; Gaps 22;
 QY 58 IIAAAASWIITCVLLC-----CSKHARCFIL-----LVFL--SCGLRE- 93
 Db 3 IIAIPAWMLLCLFILSTPTTIVTSGLPSTLRCFSIDGCTQMTSTLVFLGFAVGLSL 62
 QY 94 -----GRNALIAAGTGIVLGHVENIHFNFKGLLDGWTCLNRAKSFHPLKLYIE 146
 Db 63 GRLSDIYGRRPVILLGLFIYVSSIIISF-----SPNIEMLMIARFIQ 105
 QY 147 A-----IQWYGLATPLSVF-----DDLVSNQTLA 172
 Db 106 AFGVSGSVIGQSMARDSYQGAELSYVIAIUSPMLLFPALGSYIGGYIIIEYLSWHYFI 165
 QY 173 VSLFSPSHVLEAQLNDSKEVLSVLYQMATTEVLSLQKLLAFAGLSVLVLTGCLFNMK 232
 Db 166 --PFLAGTI-----LLALYQILPETNYIAFSQSKYFEVENIIKDKMLWLY 213
 QY 233 RF-LGPGCKMYENIYITQVQFDERERHQPCVLPLNKEERRYIPIPTWPT----- 286
 Db 214 AFIIGAFNGIYVGFIEAPFILIDQMR-----VLPSPFYGLAFLL 253
 QY 287 -----PKERNKLGFLFLILHLCIWLFAA-----VDYLLYRLI 321
 Db 254 SFASIFGFLGGYLIKRRQVYDKKVMSTGTF-----SLCGCILFAVDSFILEFILVSNV 308
 QY 322 FSVSKQFOSLPGEVH-----LKLHGEKQGTODIHDSSFNISVPEPN-C 365
 Db 309 FAIAMIWPM--MIHMGHSLIAITURYALEDYATVTGTAGSIFGAIYVYVIASTYC 365
 QY 366 IPKPKFLSETWPSVLILVLMGLSSILQMLKILVSFSPVERKRIQYLHAKIL 425
 Db 366 VSK--IHGETISNLSLLCIVL-----SISSVI-----SFY-----YI-CLLY 399
 QY 426 KKRS 429
 Db 400 KKKS 403

RESULT 13
 Q9PLZ4 PRELIMINARY; PRT; 391 AA.
 AC Q9PLZ4;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DE PUTATIVE TRANSMEMBRANE TRANSPORT PROTEIN.
 GN CJI1684C.
 OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
 OC Campylobacter.
 OX NCBI_TaxID=197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCTC 11168;
 RX MEDLINE=20150912; PubMed=10688204;
 RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
 RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
 RA Jagers K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Rajandream M.A., Rutherford K.M., Van Vliet A.H.M.,
 RA Whitehead S., Barrell B.G.:
 RT "The genome sequence of the food-borne pathogen *Campylobacter jejuni*
 reveals hypervariable sequences.";
 RL Nature 403:665-668(2000).
 DR EMBL; AL139079; CAB73670.1; -.
 KW Complete proteome.
 SQ SEQUENCE 391 AA; 44004 MW; 43CE55004589DE46 CRC64;

Query Match 4.2%; Score 102; DB 2; Length 391;
 Best Local Similarity 20.9%; Pred. No. 2.5;

QY 243 ENIYTRQFVQDERERHQORPCVLPNKEERRK-----YVITPTFWPTPK 289
Db 73 --IYFNFFVK--KRSQDLGVLAIGFSKRELTLLTLENLVILVLSYLSLLGPT--- 125
QY 290 RKNLGLFELPILI--HLC-----IWLFAVDYLLYRLIFSFSVKQ 327
Db 126 -----LYFLAVLVITHLLDLMTEVQWFIIVKELIESLGLVWVFLINFTNGVI--IAK- 177
QY 328 FQSLPGFEVHLKLHGEK----OQTODIHDSSFNIS-----VFEPNCIPKPKFLLSET 376
Db 178 -QSLIEF-VNFSKKAEEKIRKVRATIAITALLSYVLCITTVFSST----RKMLLSVG 231
QY 377 WVPLSVILLILVMGLLSSILMOLKILVSASFYPSVERKRIQVLAHAKLLKRSKQPLGEV 436
Db 232 MVPISLIIILVILGSIFTIRYGLTFVIS-----FLKEK--KKRLYRPLSNI 276

Search completed: January 31, 2002, 08:09:15
Job time: 446 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2002, 08:00:39 ; Search time 56.8 Seconds
(without alignments)
612.930 Million cell updates/sec

Title: US-09-713-098-2
Perfect score: 2448
Sequence: 1 MGWTSFTDIFLSWEIVVS.....LPVLKMRKKONDMASADKS 470

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_1101.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	2448	100.0	470	22 AAE02638 Human dendritic ce
2	1799.5	73.5	352	22 AAB87357 Human gene 16' enco
3	1314	53.7	292	21 AAB34797 Human secreted pro
4	1309	53.5	257	22 AAE03917 Human gene 20' enco
5	1309	53.5	257	22 AAB87399 Human gene 16' enco
6	1309	53.5	257	22 AAB87400 Human gene 16' enco
7	1309	53.5	257	22 AAB87401 Human gene 16' enco
8	1303	53.2	291	22 AAB87401 Human gene 16' enco
9	1171	7.0	485	20 AAY45267 Human secreted pro
10	1116	4.7	285	21 AAG55749 Arabidopsis thalia
11	1116	4.7	315	21 AAG55748 Arabidopsis thalia

12	113	4.6	285	21	AAG43662	Arabidopsis thalia
13	113	4.6	315	21	AAG43661	Arabidopsis thalia
14	101	4.1	681	21	AA54070	Enzyme EPS6 which
15	101	4.1	681	21	AA543772	Anino acid sequenc
16	99	4.0	224	21	AAB34709	Human secreted pro
17	99	4.0	242	22	AAM39089	Human polypeptide
18	99	4.0	242	22	AAB88465	Human membrane or
19	99	4.0	263	22	AA572603	Human Electron Tra
20	99	4.0	687	21	AA569165	A human N-acetylgl
21	99	4.0	765	21	AA569167	A mature human N-a
22	99	4.0	765	21	AA569168	A mature human N-a
23	99	4.0	997	22	AAB90764	Human shear stress
24	99	4.0	1485	21	AA569166	A mature human N-a
25	96	3.9	405	19	AA568502	H. pylori GPO 119
26	94.5	3.9	196	21	AAG43663	Arabidopsis thalia
27	94.5	3.9	196	21	AAG55750	Arabidopsis thalia
28	93.5	3.8	614	21	AAG51000	Arabidopsis thalia
29	93	3.8	337	18	AAW09110	Human amine recept
30	93	3.8	337	20	AA513736	Human amine recept
31	93	3.8	391	20	AA534946	Chlamydia pneumoni
32	92	3.8	447	22	AA572393	Rice LH1 encoded
33	91	3.7	246	20	AA534805	Chlamydia pneumoni
34	91	3.7	312	22	AAG72748	Human olfactory re
35	91	3.7	400	22	AAU03803	G protein-coupled
36	91	3.7	468	12	AA521019	sod2 gene-encoded
37	91	3.7	1498	16	AA577085	Hamster sulphonylu
38	91	3.7	1582	16	AA577088	Hamster sulphonylu
39	91	3.7	2132	18	AAW21739	Variant rat DRG (S
40	90.5	3.7	489	21	AAB36477	Human SNORF36b rec
41	90	3.7	215	16	AA575908	Human olfactory re
42	90	3.7	309	22	AAB87785	Rat T2R05 amino ac
43	90	3.7	310	22	AAG72727	Human olfactory re
44	90	3.7	312	22	AAG72289	Human olfactory re
45	90	3.7	473	22	AAG81483	S. epidermidis ope

ALIGNMENTS

RESULT 1
AAE02638
ID AAE02638 standard; Protein; 470 AA.
XX AC AAE02638;
XX AC
DT 06-AUG-2001 (first entry)
XX
DE Human dendritic cell specific transmembrane protein (DC-STAMP).
XX
KW Human; dendritic cell specific transmembrane protein; DC-STAMP;
KW forensic science; therapy; abnormal physiology; allergic condition;
KW asthma; cancer; autoimmune disease; diabetes mellitus;
KW drug screening.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 35..51
FT /label= Transmembrane_domain_1
FT Domain 57..75
FT /label= Transmembrane_domain_2
FT Domain 96..114
FT /label= Transmembrane_domain_3
FT Domain 144..162
FT /label= Transmembrane_domain_4
FT Modified-site 168..170
FT /note= "Asn is N-glycosylated"
FT Modified-site 187..189
FT /note= "Asn is N-glycosylated"
FT Domain 214..230
FT /label= Transmembrane_domain_5
FT Domain 295..313
FT /label= Transmembrane_domain_6

CC disorders, and infections. The proteins can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin aging due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues, to identify their
CC cognate ligands or binding partners, and in chemotaxis, and can be used
CC as a food additive or preservative to modify storage properties.
CC Antibodies specific for a protein of the invention can be used in
CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunosorbent assay (ELISA). The present sequence represents a human
CC secreted protein of the invention.
XX
SQ Sequence 352 AA;

Query Match 73.5%; Score 1799.5; DB 22; Length 352;
Best Local Similarity 96.4%; Pred. No. 6.4e-180;
Matches 344; Conservative 2; Mismatches 4; Indels 7; Gaps 1;

QY 27 MDFIQHLGVCCCLVALISVGLLSVAACWFLPSITIAAASWIITCVLLCCSKHARCFILLVF 86
Db 1 mdfiqhlgvccclvalisvglsvaaacwflpsiaaaswiitcvllccskharcfillvf 60

QY 87 LSCGLREGRNALIAAGTGVILGHVENIFHNFKGLLDGTCNLRKSFSTHFPLKKYIE 146
Db 61 lscglregrnaliaagtgvilghvenifhnfkglldgmcnlraksfshfplkkyle 120

QY 147 AIQWIYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTVE 206
Db 121 aiqwiyglatplsvfddlvswnqtlavslfshvleaqndskgevlsvlyqmattev 180

QY 207 LSSLGKQLAFAGLSVLVLTGLFMKRFGLPCGWKYENIYTRQVQFDERERHQRPCV 266
Db 181 lsslqkllafaglsvlvltgfmkrfglpcgwkyniytrqvqfdererhqrpcv 240

QY 267 LPLNKEERKYYIIPTFWPTPKRNKLGFLFPLILHLICWVLFVAAVDYLLYLIFS 326
Db 241 lplnkeerryiipftwptpkrnknlglfplilhlciwlvfaavdyllylifsvsk 300

QY 327 QFOSLPGFEVHLKLGKEGQTDIIHDSFNSVFEPNCPKPKFLLSETWPLSVI 383
Db 301 qfoslpgfevhlkhlgekgtdiindhssfnsvfepncipkpkfllsetwplsvi 350

RESULT 3
AAB34797
ID AAB34797 standard; Protein; 292 AA.
XX AAB34797;
AC AAB34797;
DT 26-JAN-2001 (first entry)
XX
DE Human secreted protein sequence encoded by gene 25 SEQ ID NO:85.
XX
KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
KW nootropic; neuroprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiant; gene therapy;
KW cancer; immune disorder; cardiovascular disorder; wound healing;
KW neurological disease; infectious disease; chromosome identification.
OS Homo sapiens.
XX
XX WO200058356-A1.
XX
PD 05-OCT-2000.
XX
PF 22-MAR-2000; 2000WO-US07535.
XX
PR 26-MAR-1999; 99US-0126511.
PR 17-DEC-1999; 99US-0172413.
XX
PA (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Komatsoulis G;
XX WPI; 2000-594639/56.
DR N-PSDB; AAC59990.
XX
PT Fifty nucleic acid molecules encoding human secreted proteins, useful
PT in the prevention, treatment and diagnosis of cancer, immune disorders,
PT cardiovascular disorders and neurological diseases -
XX
PS Claim 1; Page 385-386; 425pp; English.
XX
CC The polynucleotide sequences given in AAC59966 to AAC60015 encode the
CC human secreted proteins given in AAB34773 to AAB34822. AAB34823 to
CC AAB34852 represent human secreted polypeptide sequences and proteins
CC homologous to them, which are given in the exemplification of the present
CC invention. Human secreted proteins have activities based on the tissues
CC and cells the genes are expressed in. Examples of activities include:
CC cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
CC antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
CC vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic; and
CC cardiant. The polynucleotides and polypeptides are useful for
CC preventing, treating or ameliorating a medical condition in e.g. humans,
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. The polynucleotides are
CC useful for chromosome identification. They are also useful as probes for
CC diagnosing a disorder related to the female reproductive system,
CC particularly breast and/or ovary cancer. They are also useful in the gene
CC therapy of breast and ovarian cancer. Nucleic acids, protein, antibodies,
CC agonists and antagonists from the present invention are useful in the
CC diagnosis, treatment and prevention of cancer, immune disorders,
CC cardiovascular disorders, wound healing, neurological diseases and
CC infectious disease. AAC59957 to AAC59965 and AAB34772 represents sequence
CC used in the exemplification of the present invention.
XX
SQ Sequence 292 AA;

Query Match 53.7%; Score 1314; DB 21; Length 292;
Best Local Similarity 92.5%; Pred. No. 3.6e-129;
Matches 259; Conservative 4; Mismatches 5; Indels 12; Gaps 3;

QY 27 MDFIQHLGVCCCLVALISVGLLSVAACWFLPSITIAAASWIITCVLLCCSKHARCFILLVF 86
Db 1 mdfiqhlgvccclvalisvglsvaaacwflpsiaaaswiitcvllccskharcfillvf 60

QY 87 LSCGLREGRNALIAAGTGVILGHVENIFHNFKGLLDGTCNLRKSFSTHFPLKKYIE 146
Db 61 lscglregrnaliaagtgvilghvenifhnfkglldgmcnlraksfshfplkkyle 120

QY 147 AIQWIYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTVE 206
Db 121 aiqwiyglatplsvfddlvswnqtlavslfshvleaqndskgevlsvlyqmattev 180

QY 207 LSSLGKQLAFAGLSVLVLTGLFMKRFGLPCGWKYENIYTRQVQFDERERHQRPCV 266
Db 181 lsslqkllafaglsvlvltgfmkrfglpcgwkyniytrqvqfdererhqrpcm 240

QY 267 LPLNKEERKYYIIPTFWPTPKRNKLGFLFPLILHLICI 306
Db 241 lplnkeerryiipftwptpkrnknlglfplilhlci 268

RESULT 4
AAE03917
ID AAE03917 standard; Protein; 257 AA.
XX
AC AAE03917;
XX
DT 09-AUG-2001 (first entry)
XX
DE Human gene 20 encoded secreted protein HMADJ14, SEQ ID NO:80.

XX Human: secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angioinetic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
KW endocrine disorder; infection; wound healing; vulnery;
KW cell culture; chemotaxis; food additive; gene therapy;
KW binding partner identification.
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
FT Peptide 1..26
FT /label= signal_peptide
FT Protein 27..257
FT /note= "Mature secreted protein"
XX W0200077022-A1.
XX
XX 21-DEC-2000.
XX
XX 01-JUN-2000; 2000WO-US15136.
XX
XX 11-JUN-1999; 99US-0138629.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Komatsoulis GA;
XX
XX WPI; 2001-367020/38.
XX N-PSDB; AAD08364.
XX
XX Nucleic acids encoding 50 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating diseases, e.g. Parkinson's
PT disease, botulism, cancers and Sclimtar syndrome -
XX
XX Claim 11; Page 535-536; 614pp; English.
XX
XX AAD08345-AAD08394 represent cDNAs corresponding to 50 human secreted
CC protein genes and AAE03898-AAE03947 represent the proteins they encode.
CC AAE03948-AAE03996 represent human secreted protein fragments or variants.
CC The genes and their secreted proteins are useful for preventing,
CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Pathological conditions can be diagnosed by determining the
CC amount of the new protein in a sample or by determining the presence of
CC mutations in the new genes. Specific uses are described for each of the
CC 50 genes, based on the tissues in which they are most highly expressed,
CC and include developing products for the diagnosis or treatment of
CC proliferative disorders, cancer, tumours, foetal and developmental
CC abnormalities, haematopoietic disorders, diseases of the immune system,
CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
CC allergies, neurological disorders (e.g., Alzheimer's disease,
CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC cardiovascular disorders, angioinetic disorders, kidney disorders,
CC gastrointestinal disorders, pregnancy-related disorders, endocrine
CC disorders, and infections. The proteins can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin aging due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues, to identify their
CC cognate ligands or binding partners, and in chemotaxis, and can be used
CC as a food additive or preservative to modify storage properties.
CC Antibodies specific for a protein of the invention can be used in
CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunosorbent assay (ELISA). The present sequence represents a human
CC secreted protein of the invention.
XX
XX Sequence 257 AA;

Query Match 53.5%; Score 1309; DB 22; Length 257;
Best Local Similarity 99.2%; Pred. No. 1e-128;
Matches 250; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 27 MDTQHLGVCCLVALLSVGLLSVAACWFLPSIIAASWITCVLLCCSKHARCFILLVF 86
Dy 1 mdfqhlgvccclvalisvllsvaaacwflpsiaaaswiltcvllccskharcfillvf 60
Qy 87 LSCGLREGRNALIAAGTGIVILGHVENIFHNFKGLDGMTCNLRKSFSTHFPPLKKYIE 146
Dy 61 lscglregrnaliaagtgivilghvenifhnmfkgllidgntcnlraksfshfpllkxyie 120
Qy 147 AIQWYIGLATPLSVFDDLVSWNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTVEV 206
Dy 121 aiqwyiglatplsvfddlvswngtllavslfspshvleaglnndskgevlsvlyqmattev 180
Qy 207 LSSIQOKLLAFAGLSLVLLGTGLFMKRFGLPGCWKYENIYTRQFVOFDERHRHQPCV 266
Dy 181 lssiqokllafaglsllvllgtglfmkrflgpcgkyenyitrqfvqfdererhqrpcv 240
Qy 267 LPLNKEERRKYV 278
Dy 241 lplnkeerrkfi 252
RESULT 5
AAB87399
ID AAB87399 standard; Protein; 257 AA.
XX AC AAB87399;
XX
XX 22-MAY-2001 (first entry)
XX
XX Human gene 16 encoded secreted protein HMADJ14, SEQ ID NO:140.
XX
XX Human; secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angioinetic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
KW endocrine disorder; infection; wound healing; vulnery;
KW cell culture; chemotaxis; food additive;
KW binding partner identification.
XX
XX Homo sapiens.
XX
XX W0200118022-A1.
XX
XX 15-MAR-2001.
XX
XX 31-AUG-2000; 2000WO-US24008.
XX
XX 03-SEP-1999; 99US-0152315.
XX 03-SEP-1999; 99US-0152317.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ni J, Baker KP, Birse CE, Fiscella M, Komatsoulis GA, Rosen CA;
PI Soppet DR, Young PE, Ebner R, Duan DR, Olsen HS, Lafleur DW;
PI Moore PA, Shi Y, Wei Y, Florence KA;
XX
XX WPI; 2001-203081/20.
XX N-PSDB; AAF91915.
XX
XX Nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT Parkinson's diseases and cancers -
XX

PS Claim 11; Page 567-568; 607pp; English.

XX AAF91858-AAF91929 represent cDNAs corresponding to 52 human secreted

CC protein genes, and AAB87342-AAB87413 represent the proteins they encode.

CC AAB87414-AAB87454 represent human secreted protein fragments. The genes

CC and their corresponding secreted proteins are useful for preventing,

CC treating or ameliorating medical conditions, e.g., by protein or gene

CC therapy. Pathological conditions can be diagnosed by determining the

CC amount of the new protein in a sample or by determining the presence of

CC mutations in the new genes. Specific uses are described for each of the

CC 52 genes, based on the tissues in which they are most highly expressed,

CC and include developing products for the diagnosis or treatment of

CC proliferative disorders, cancer, tumours, foetal and developmental

CC abnormalities, haematopoietic disorders, diseases of the immune system,

CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,

CC allergies, neurological disorders (e.g., Alzheimer's disease,

CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,

CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,

CC cardiovascular disorders, angiogenic disorders, kidney disorders,

CC gastrointestinal disorders, pregnancy-related disorders, endocrine

CC disorders, and infections. The proteins can also be used to aid wound

CC healing and epithelial cell proliferation, to prevent skin aging due to

CC sunburn, to maintain organs before transplantation, for supporting cell

CC culture of primary tissues, to regenerate tissues, to identify their

CC cognate ligands or binding partners, and in chemotaxis, and can be used

CC as a food additive or preservative to modify storage properties.

CC Antibodies specific for a protein of the invention can be used in

CC alleviating symptoms associated with the disorders mentioned above, and

CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked

CC immunosorbent assay (ELISA). The present sequence represents a human

CC secreted protein of the invention.

XX Sequence 257 AA;

SEQ

Query Match 53.5%; Score 1309; DB 22; Length 257;

Best Local Similarity 99.2%; Pred. No. 1e-128; 0; Indels 0; Gaps 0;

Matches 250; Conservative 2; Mismatches 0;

QY 27 MDFIQLGVCCCLVALISVGLLSVAACWFLPSIIAAASWIITCVLLCCSKHRCFILLVF 86

DB 1 mdfiqhgvccclvalisvglsvaacwflpsiaaaswiitcvllccskharcfllvlf 60

QY 87 LSCGLREGRNALTAAGTGVILGHVENIFNFKGLDGMTCNIRAKSFSTHFPLKKYIE 146

DB 61 lscglregrnaliaagtgvilghvenifnfgkllidgmcniraksfshfpllkyie 120

QY 147 AIOWIYGLATPLSVFDLYSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 206

DB 121 alqwiylglatplsvfdlysnqtlavslfspshvleaqindskgevislyqmattev 180

QY 207 LSSLGOKLLAFAGLSVLLTGTGLFMRFLGPCGKWKYENIVITRQFVQFDERERHQQRPCV 266

DB 181 lsslgokllafaglsvlltgtglfmrflgpcgwkkyenyitrfqvqfdererhqqrpcv 240

QY 267 LPUNKERRRYV 278

DB 241 lpinkeerrkfi 252

RESULT 6

AAB87400

ID AAB87400 standard; Protein; 257 AA.

XX AAB87400;

XX 22-MAY-2001 (first entry)

XX Human gene 16 encoded secreted protein HMADJ14, SEQ ID NO:141.

DE Human; secreted protein; proliferative disorder; cancer; tumour;

XX foetal abnormality; developmental abnormality; haematopoietic disorder;

KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;

KW

KW inflammation; allergy; neurological disorder; Alzheimer's disease;

KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;

KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;

KW cardiovascular disorder; angiogenic disorder; kidney disorder;

KW gastrointestinal disorder; pregnancy-related disorder;

KW endocrine disorder; infection; wound healing; vulnerary;

KW cell culture; chemotaxis; food additive;

KW binding partner identification.

XX Homo sapiens.

OS WO200118022-A1.

XX 15-MAR-2001.

PD 31-AUG-2000; 2000WO-US24008.

XX 03-SEP-1999; 99US-0152315.

XX 03-SEP-1999; 99US-0152317.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ni J, Baker KP, Birse CE, Fiscella M, Komatsoulis GA, Rosen CA;

XX Soppet DR, Young PE, Ebner R, Duan DR, Olsen HS, Lafleur DW;

PI Moore PA, Shi Y, Wei Y, Florence KA;

XX WPI; 2001-203081/20.

DR N-PSDB; AAF91916.

XX Nucleic acid molecules encoding human secreted proteins, used in

XX preventing, treating or ameliorating a disorder, e.g. Alzheimer's and

XX Parkinson's diseases and cancers -

XX Claim 11; Page 568-569; 607pp; English.

XX AAF91858-AAF91929 represent cDNAs corresponding to 52 human secreted

CC protein genes, and AAB87342-AAB87413 represent the proteins they encode.

CC AAB87414-AAB87454 represent human secreted protein fragments. The genes

CC and their corresponding secreted proteins are useful for preventing,

CC treating or ameliorating medical conditions, e.g., by protein or gene

CC therapy. Pathological conditions can be diagnosed by determining the

CC amount of the new protein in a sample or by determining the presence of

CC mutations in the new genes. Specific uses are described for each of the

CC 52 genes, based on the tissues in which they are most highly expressed,

CC and include developing products for the diagnosis or treatment of

CC proliferative disorders, cancer, tumours, foetal and developmental

CC abnormalities, haematopoietic disorders, diseases of the immune system,

CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,

CC allergies, neurological disorders (e.g., Alzheimer's disease,

CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,

CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,

CC cardiovascular disorders, angiogenic disorders, kidney disorders,

CC gastrointestinal disorders, pregnancy-related disorders, endocrine

CC disorders, and infections. The proteins can also be used to aid wound

CC healing and epithelial cell proliferation, to prevent skin aging due to

CC sunburn, to maintain organs before transplantation, for supporting cell

CC culture of primary tissues, to regenerate tissues, to identify their

CC cognate ligands or binding partners, and in chemotaxis, and can be used

CC as a food additive or preservative to modify storage properties.

CC Antibodies specific for a protein of the invention can be used in

CC alleviating symptoms associated with the disorders mentioned above, and

CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked

CC immunosorbent assay (ELISA). The present sequence represents a human

CC secreted protein of the invention.

XX Sequence 257 AA;

SEQ

Query Match 53.5%; Score 1309; DB 22; Length 257;

Best Local Similarity 99.2%; Pred. No. 1e-128; 0; Indels 0; Gaps 0;

Matches 250; Conservative 2; Mismatches 0;

QY 27 MDFIQLGVCCCLVALISVGLLSVAACWFLPSIIAAASWIITCVLLCCSKHRCFILLVF 86

Db 1 mdfihlgvccvalisvlgslsvaacwflpsiaaaswiitcvllccskharcfillvlf 60
Qy 87 LSCGLREGRNALIAAGTGIVILGHVENIFHNFKGLLDGMCNLRKAKSFIHFPLKKYIE 146
Db 61 lscglregrnallaaagtgivilghvenifhnfkglldgmcntraksfsihfplkkyle 120
Qy 147 AIQWYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 206
Db 121 aiqwylglatplsvfddlvsnqtlavslfspshvleaqlnndskgevlsvlyqmattev 180
Qy 207 LSSLGOKLLAFAGLSVLGTLGFLPMKRFGLGPCGKWKYENIYITROFVQDERHQRPCV 266
Db 181 lsslgokllafaglsvlgtlglfkmkrflgpcgwkyniyitrqfvqderhqrpcv 240
Qy 267 LPLNKEERRKYV 278
Db 241 lplnkeerrkfi 252
RESULT 7
AAB87454
ID AAB87454 standard; Protein: 257 AA.
XX
AC AAB87454;
XX
DT 22-MAY-2001 (first entry)
XX
DE Human gene 16 encoded secreted protein fragment.
XX
KW Human; secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angioinetic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
KW endocrine disorder; infection; wound healing; vulnery;
KW cell culture; chemotaxis; food additive;
KW binding partner identification.
XX
OS Homo sapiens.
XX
PF WO200118022-A1.
XX
PN 31-AUG-2000; 2000WO-US24008.
XX
PD 15-MAR-2001.
XX
PF 03-SEP-1999; 99US-0152315.
XX
PR 03-SEP-1999; 99US-0152317.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ni J, Baker KP, Birse CE, Fiscella M, Komatsoulis GA, Rosen CA;
PI Soppet DR, Young PE, Ebner R, Duan DR, Olsen HS, Lafleur DW;
PI Moore PA, Shi Y, Wei Y, Florence KA;
XX
DR WPI: 2001-203081/20.
XX
XX Nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT Parkinson's diseases and cancers -
XX
PS Disclosure; Page 55; 607pp; English.
XX
CC AAF91858-AAF91929 represent cDNAs corresponding to 52 human secreted
CC protein genes, and AAB87342-AAB87413 represent the proteins they encode.
CC AAB87414-AAB87454 represent human secreted protein fragments. The genes
CC and their corresponding secreted proteins are useful for preventing,
CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Pathological conditions can be diagnosed by determining the

CC amount of the new protein in a sample or by determining the presence of
CC mutations in the new genes. Specific uses are described for each of the
CC 52 genes, based on the tissues in which they are most highly expressed,
CC and include developing products for the diagnosis or treatment of
CC proliferative disorders, cancer, tumours, foetal and developmental
CC abnormalities, haematopoietic disorders, diseases of the immune system,
CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
CC allergies, neurological disorders (e.g., Alzheimer's disease,
CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC cardiovascular disorders, angioinetic disorders, kidney disorders,
CC gastrointestinal disorders, pregnancy-related disorders, endocrine
CC disorders, and infections. The proteins can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin aging due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues, to identify their
CC cognate ligands or binding partners, and in chemotaxis, and can be used
CC as a food additive or preservative to modify storage properties.
CC Antibodies specific for a protein of the invention can be used in
CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunosorbent assay (ELISA). The present sequence represents a human
CC secreted protein fragment referred to in the disclosure of the invention.
XX
SQ Sequence 257 AA;
Query Match 53.5%; Score 1309; DB 22; Length 257;
Best Local Similarity 99.2%; Pred. No. 1e-128; 0; Indels 0; Gaps 0;
Matches 250; Conservative 2; Mismatches 0;
Qy 27 MDFIHLGVCCLVALISVGLSLVAACWFLPSIIAAAASWIITCVLLCCSKHARCFILLVF 86
Db 1 mdfihlgvccvalisvlgslsvaacwflpsiaaaswiitcvllccskharcfillvlf 60
Qy 87 LSCGLREGRNALIAAGTGIVILGHVENIFHNFKGLLDGMCNLRKAKSFIHFPLKKYIE 146
Db 61 lscglregrnallaaagtgivilghvenifhnfkglldgmcntraksfsihfplkkyle 120
Qy 147 AIQWYGLATPLSVFDDLVSNQTLAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTTEV 206
Db 121 aiqwylglatplsvfddlvsnqtlavslfspshvleaqlnndskgevlsvlyqmattev 180
Qy 207 LSSLGOKLLAFAGLSVLGTLGFLPMKRFGLGPCGKWKYENIYITROFVQDERHQRPCV 266
Db 181 lsslgokllafaglsvlgtlglfkmkrflgpcgwkyniyitrqfvqderhqrpcv 240
Qy 267 LPLNKEERRKYV 278
Db 241 lplnkeerrkfi 252
RESULT 8
AAB87401
ID AAB87401 standard; Protein: 291 AA.
XX
AC AAB87401;
XX
DT 22-MAY-2001 (first entry)
XX
DE Human gene 16 encoded secreted protein HMADJ74, SEQ ID NO:142.
XX
KW Human; secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angioinetic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
KW endocrine disorder; infection; wound healing; vulnery;
KW cell culture; chemotaxis; food additive;
KW binding partner identification.

XX Homo sapiens.
XX WO200118022-A1.
XX 15-MAR-2001.
XX 31-AUG-2000; 2000WO-US24008.
XX 03-SEP-1999; 99US-0152315.
XX 03-SEP-1999; 99US-0152317.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Ni J, Baker KP, Birse CE, Fiscella M, Komatsoulis GA, Rosen CA;
PI Soppet DR, Young PE, Ebner R, Duan DR, Olsen HS, Lafleur DW;
PI Moore PA, Shi Y, Wei Y, Florence KA;
XX WPI: 2001-203081/20.
XX N-PSDB; AAF91917.
XX Nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT Parkinson's diseases and cancers -
XX Claim 11; Page 569-570; 607pp; English.
XX AAF91858-AAF91929 represent cDNAs corresponding to 52 human secreted
CC protein genes, and AAB87342-AAB87413 represent the proteins they encode.
CC AAB87414-AAB87454 represent human secreted protein fragments. The genes
CC and their corresponding secreted proteins are useful for preventing,
CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Pathological conditions can be diagnosed by determining the
CC amount of the new protein in a sample or by determining the presence of
CC mutations in the new genes. Specific uses are described for each of the
CC 52 genes, based on the tissues in which they are most highly expressed,
CC and include developing products for the diagnosis or treatment of
CC proliferative disorders, cancer, tumours, foetal and developmental
CC abnormalities, haematopoietic disorders, diseases of the immune system,
CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
CC allergies, neurological disorders (e.g., Alzheimer's disease,
CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC cardiovascular disorders, angiogenic disorders, kidney disorders,
CC gastrointestinal disorders, pregnancy-related disorders, endocrine
CC disorders, and infections. The proteins can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin aging due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues, to identify their
CC cognate ligands or binding partners, and in chemotaxis, and can be used
CC as a food additive or preservative to modify storage properties.
CC Antibodies specific for a protein of the invention can be used in
CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunosorbent assay (ELISA). The present sequence represents a human
XX secreted protein of the invention.
XX Sequence 291 AA;
SQ
Query Match 53.2%; Score 1303; DB 22; Length 291;
Best Local Similarity 92.1%; Pred. No. 5.1e-128;
Matches 258; Conservative 4; Mismatches 6; Indels 12; Gaps 3;
QY 27 MFIQHLGVCCILVALISVGLLSVAACWFLPSIIAAASWIIITVLCCSKHRCFILLVF 86
DB 1 mdfighlgvccilvalisvlgllsvaacwflpsiaaaswiitcvllccskhrcfllvf 60
QY 87 LSCGLREGNRLIAAGTGVILGHVENIFHNFKGLDGMTCNLRKSFSTHFPPLKKYIE 146
DB 61 lxsgrregnrliaagtgivilghvenifhnrkglldgmtnlrksfsthfpplkkyle 120
QY 147 AIQWIYGLATPLSVFDDLSVWQTLAVSLFSPSHVLEAQLNDSKGEVLSVLVQMATTTEV 206

Db 121 aiqwiylatplsvfddlsvwnqtlavslfshvleaqldnsgkgevlsvlyqmatteev 180
QY 207 LSSLGQKLLAFAGLSVLVLLGTGLFMKRLPGCGWKYENIYITRQVFQFDERERHQRPCV 266
Db 181 lsslqgkllafaglsvlvllgtglfmkrflpgcgwkyniyltrqvfqfdererhqrpcv 240
QY 267 LPLNKEERKYYIITPFWPTPKERKNLGLFPLPILHLICI 306
Db 241 lplnkeerrk-----nkelkilm-llp-lyici 268
RESULT 9
AA45267
ID AAY45267 standard; Protein; 485 AA.
XX AC
XX AAY45267;
XX 23-NOV-1999 (first entry)
XX Human secreted protein encoded from gene 11.
XX Secreted protein; human; gene therapy; diagnosis; treatment; cancer;
KW protein therapy; tumor; neurodegenerative disorder; blood disorder; AIDS;
KW developmental abnormality; leukemia; immune system; autoimmune disease;
KW hepatic disease; renal disease; inflammation; allergy; schizophrenia;
KW Alzheimer's disease; cognitive disorder; arthritis; infection; psoriasis;
KW transplant rejection; diabetes; asthma; sepsis; acne; metabolic disorder;
KW cardiovascular disorder; food additive; preservative.
XX Homo sapiens.
XX WO9946289-A1.
XX 16-SEP-1999.
XX 11-MAR-1999; 99WO-US05721.
XX 12-MAR-1998; 98US-0077686.
PR 12-MAR-1998; 98US-0077687.
PR 12-MAR-1998; 98US-0077696.
PR 12-MAR-1998; 98US-0077714.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Ruben SM, Ferrie AM, Rosen CA, Florence C, Young PE, Yu G, Ni J;
DR WPI: 1999-551363/46.
DR N-PSDB; AAZ27243.
XX New isolated human genes, useful for diagnosis and treatment of, e.g.
XX cancers -
XX Claim 1b; 179-180; 306pp; English.
XX This invention describes novel human genes and the secreted proteins
CC they encode. The polynucleotides and their corresponding secreted
CC polypeptides are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also pathological conditions
CC can be diagnosed by determining the amount of the new polypeptides in a
CC sample or by determining the presence of mutations in the new
CC polynucleotides. Specific uses are described for each of the
CC polynucleotides of the invention, based on which tissues they are most
CC highly expressed in, and include developing products for the diagnosis or
CC treatment of cancer, tumors, neurodegenerative disorders, developmental
CC abnormalities, blood disorders, leukemias, diseases of the immune system,
CC autoimmune diseases, hepatic and renal disease, inflammation, allergies,
CC Alzheimer's and cognitive disorders, schizophrenia, arthritis,
CC infections, AIDS, transplant rejection, diabetes, asthma, sepsis, acne,
CC psoriasis, cardiovascular disorders, and metabolic disorders. The
CC polypeptides or polynucleotides can also be used as food additives or
CC preservatives. The polypeptides are also useful for identifying their
CC binding partners. AAY45267-Y45289 represent the human secreted proteins

CC described in the method of the invention.

XX Sequence 485 AA;

Query Match 7.0%; Score 171; DB 20; Length 485;
Best Local Similarity 20.6%; Pred. No. 4.6e-09;
Matches 94; Conservative 63; Mismatches 167; Indels 132; Gaps 18;

QY 35 VCVLVALISVGLLSVAACWFLPSPHIAAASWIITCVLLCCSKHARCFILVFLSCGLREG 94
 Db 5 vcglvflslgl-----vppv-----rcifalsvptlmeqg 36
 QY 95 RNALIAAGTGIVILGHVENIFHNFKGLLDGTCNLRKSPSIHPLKXIEATQWYGL 154
 Db 37 rrillsystatlaavpvnlanvgaagqlrcvctegs-----lesllntthqihaa 88
 QY 155 ATPLSVFDLVSNNOTLAVSLFSPSHVLEAQLNDSKGEVLSVLQMATTEVLSLG--Q 212
 Db 89 sralgptgqagsrglt-----feaqdnsgs-----afyllhmltvtqgvldefsgle 133
 QY 213 KLLAFAGLSVLVLTGTLGFMKRFGLPCGW-----KYENIYITROFVQ--FDERERHQ 261
 Db 134 slaraaalgtrvvgtlglmlglivesawylhcyldlrfdniyatqltqlrlaqaqathl 193
 QY 262 ORP-----CWLPLNKEERRKYVIPTFWPTPKERNLGLFELPILIHLCIHWLFAAV 313
 Db 194 lappptwllgaagqlrisqeellscell-----rlgllal-llvatavav---at 237
 QY 314 DYLLRLIFSVMQFQSLPGFEVHLKLGKQGTQDIIHSSFNISVFEP-----NCIPKP 369
 Db 238 dhvafilaqatvdwaqkltptvi-----tltkydvayvtlglfpflfnqlapes 287
 QY 370 KFL---LSETW-----VPLSV--ILLILVMLGLSSILMQLKTLV 404
 Db 288 pflsvhsyqweirltsarcpilparpraaxagqlagstvllegayrrlxai 347
 QY 405 SASFYSVEKRQYQYHLAKLKRKSQ-----PLGE 435
 Db 348 aasftaagearrirhlharlqrrhdxqgqqlpgd 383

RESULT 10

AAG55749

ID AAG55749 standard; Protein; 285 AA.

XX

AC AAG55749;

XX

DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 71534.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.

PR

PR 03-MAR-1999; 99US-0123180.

PR

PR 09-MAR-1999; 99US-0123348.

PR

PR 23-MAR-1999; 99US-0125788.

PR

PR 25-MAR-1999; 99US-0126264.

PR

PR 29-MAR-1999; 99US-0126785.

PR

PR 01-APR-1999; 99US-0127462.

PR

PR 06-APR-1999; 99US-0128234.

PR

PR 08-APR-1999; 99US-0128714.

PR

PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.

PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-0301439.	PR	06-JUL-1999; 99US-0142390.
XX		PR	08-JUL-1999; 99US-0142803.
XX	25-FEB-1999; 99US-0121825.	PR	09-JUL-1999; 99US-0142920.
PR	05-MAR-1999; 99US-0123180.	PR	12-JUL-1999; 99US-0142977.
PR	09-MAR-1999; 99US-0123548.	PR	13-JUL-1999; 99US-0143542.
PR	23-MAR-1999; 99US-0125788.	PR	14-JUL-1999; 99US-0143624.
PR	25-MAR-1999; 99US-0126264.	PR	15-JUL-1999; 99US-0144005.
PR	29-MAR-1999; 99US-0126785.	PR	16-JUL-1999; 99US-0144085.
PR	01-APR-1999; 99US-0127462.	PR	19-JUL-1999; 99US-0144325.
PR	06-APR-1999; 99US-0128234.	PR	19-JUL-1999; 99US-0144331.
PR	08-APR-1999; 99US-0128714.	PR	19-JUL-1999; 99US-0144332.
PR	16-APR-1999; 99US-0129845.	PR	19-JUL-1999; 99US-0144333.
PR	19-APR-1999; 99US-0130077.	PR	19-JUL-1999; 99US-0144334.
PR	21-APR-1999; 99US-0130449.	PR	19-JUL-1999; 99US-0144335.
PR	23-APR-1999; 99US-0130510.	PR	20-JUL-1999; 99US-0144352.
PR	28-APR-1999; 99US-0130891.	PR	20-JUL-1999; 99US-0144632.
PR	30-APR-1999; 99US-0131449.	PR	20-JUL-1999; 99US-0144884.
PR	30-APR-1999; 99US-0132048.	PR	21-JUL-1999; 99US-0144814.
PR	04-MAY-1999; 99US-0132407.	PR	21-JUL-1999; 99US-0145086.
PR	05-MAY-1999; 99US-0132484.	PR	21-JUL-1999; 99US-0145088.
PR	06-MAY-1999; 99US-0132485.	PR	22-JUL-1999; 99US-0145085.
PR	08-MAY-1999; 99US-0132486.	PR	22-JUL-1999; 99US-0145087.
PR	07-MAY-1999; 99US-0132487.	PR	22-JUL-1999; 99US-0145089.
PR	11-MAY-1999; 99US-0132863.	PR	23-JUL-1999; 99US-0145192.
PR	14-MAY-1999; 99US-0134256.	PR	23-JUL-1999; 99US-0145145.
PR	14-MAY-1999; 99US-0134219.	PR	23-JUL-1999; 99US-0145218.
PR	14-MAY-1999; 99US-0134221.	PR	26-JUL-1999; 99US-0145224.
PR	18-MAY-1999; 99US-0134370.	PR	27-JUL-1999; 99US-0145276.
PR	19-MAY-1999; 99US-0134768.	PR	27-JUL-1999; 99US-0145913.
PR	20-MAY-1999; 99US-0134941.	PR	27-JUL-1999; 99US-0145918.
PR	21-MAY-1999; 99US-0135124.	PR	28-JUL-1999; 99US-0145919.
PR	21-MAY-1999; 99US-0135353.	PR	02-AUG-1999; 99US-0145951.
PR	24-MAY-1999; 99US-0135629.	PR	02-AUG-1999; 99US-0146386.
PR	25-MAY-1999; 99US-0136021.	PR	02-AUG-1999; 99US-0146388.
PR	27-MAY-1999; 99US-0136392.	PR	03-AUG-1999; 99US-0146389.
PR	28-MAY-1999; 99US-0136782.	PR	03-AUG-1999; 99US-0147038.
PR	01-JUN-1999; 99US-0137222.	PR	04-AUG-1999; 99US-0147204.
PR	03-JUN-1999; 99US-0137528.	PR	04-AUG-1999; 99US-0147302.
PR	04-JUN-1999; 99US-0137502.	PR	05-AUG-1999; 99US-0147192.
PR	07-JUN-1999; 99US-0137724.	PR	05-AUG-1999; 99US-0147260.
PR	08-JUN-1999; 99US-0138094.	PR	06-AUG-1999; 99US-0147303.
PR	10-JUN-1999; 99US-0138540.	PR	06-AUG-1999; 99US-0147416.
PR	10-JUN-1999; 99US-0138847.	PR	09-AUG-1999; 99US-0147493.
PR	14-JUN-1999; 99US-0139119.	PR	09-AUG-1999; 99US-0147935.
PR	16-JUN-1999; 99US-0139452.	PR	10-AUG-1999; 99US-0148171.
PR	18-JUN-1999; 99US-0139453.	PR	11-AUG-1999; 99US-0148319.
PR	17-JUN-1999; 99US-0139492.	PR	12-AUG-1999; 99US-0148341.
PR	18-JUN-1999; 99US-0139454.	PR	13-AUG-1999; 99US-0148565.
PR	18-JUN-1999; 99US-0139455.	PR	13-AUG-1999; 99US-0148684.
PR	18-JUN-1999; 99US-0139456.	PR	16-AUG-1999; 99US-0149368.
PR	18-JUN-1999; 99US-0139457.	PR	17-AUG-1999; 99US-0149175.
PR	18-JUN-1999; 99US-0139458.	PR	18-AUG-1999; 99US-0149426.
PR	18-JUN-1999; 99US-0139459.	PR	20-AUG-1999; 99US-0149722.
PR	18-JUN-1999; 99US-0139460.	PR	20-AUG-1999; 99US-0149723.
PR	18-JUN-1999; 99US-0139461.	PR	20-AUG-1999; 99US-0149929.
PR	18-JUN-1999; 99US-0139462.	PR	23-AUG-1999; 99US-0149902.
PR	18-JUN-1999; 99US-0139463.	PR	23-AUG-1999; 99US-0149930.
PR	18-JUN-1999; 99US-0139750.	PR	25-AUG-1999; 99US-0150566.
PR	18-JUN-1999; 99US-0139763.	PR	26-AUG-1999; 99US-0150884.
PR	18-JUN-1999; 99US-0139817.	PR	27-AUG-1999; 99US-0151065.
PR	21-JUN-1999; 99US-0139817.	PR	27-AUG-1999; 99US-0151066.
PR	22-JUN-1999; 99US-0139899.	PR	27-AUG-1999; 99US-0151080.
PR	23-JUN-1999; 99US-0140353.	PR	30-AUG-1999; 99US-0151303.
PR	23-JUN-1999; 99US-0140354.	PR	31-AUG-1999; 99US-0151438.
PR	24-JUN-1999; 99US-0140695.	PR	01-SEP-1999; 99US-0151930.
PR	28-JUN-1999; 99US-0140823.	PR	07-SEP-1999; 99US-0152363.
PR	29-JUN-1999; 99US-0140991.	PR	10-SEP-1999; 99US-0153070.
PR	30-JUN-1999; 99US-0141287.	PR	13-SEP-1999; 99US-0153758.
PR	01-JUL-1999; 99US-0141842.	PR	15-SEP-1999; 99US-0154018.
PR	01-JUL-1999; 99US-0142154.	PR	16-SEP-1999; 99US-0154039.
PR	02-JUL-1999; 99US-0142055.	PR	20-SEP-1999; 99US-0154779.
		PR	22-SEP-1999; 99US-0155139.

```

PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161004.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 4.7%; Score 116; DB 21; Length 315;
Best Local Similarity 22.8%; Pred. No. 0.0014;
Matches 77; Conservative 37; Mismatches 127; Indels 96; Gaps 15;

QY 6 SGTDFL-SLMEIYVSPSPGMDPIQHLGVCLVALISVGLLSVAACWFLPSIIAAA 63
Db 53 ngsvdicddvwrpsfeqihpgwl-----lgrvlgfcfillanniarfa 96
QY 64 S-----WIITCVLLCCSKHARCFFILLVFLSC--GLREGRNALIAAGTGIVILGH 110
Db 97 nrgwriyyvtqwtftliaiyfmgms-----llsiyqclqykqgntgladqvgi----d 148
QY 111 VENIFNFKGLLDGMCNLRKASFSTHFPLKKYIEAQIWIYGLATPLSVDFDLSVSNQT 170
Db 149 aengfrs--plldgdmvsefkrktsgealksyvhlfiyqmgagavlttdsiyw--t 204
QY 171 LAVSLFSPSHLEAQLANDSKGEVLSVLYQMATTTEVLSLQKLLAFAGLSVLLGTGLF 230
Db 205 vifpfls-----lqd-----yemsfntvnlht-----snlvlllidtf 237
QY 231 MKRFLGPCWKYENIYITQFVQFDRERHQRPCVLPINKERRKYVIPTFWPTPKER 290
Db 238 lnrlkplfrfsyflwtgfcvlfq-----wil-----hmtfslvgwpypp--f 277
QY 291 KNIGLFFLPI-----LIHLCIWLFLFAAVDYLLYRLI 321
Db 278 lnlsldmapwvllvallhpsygfiallvkkykyl 314

```

RESULT 12
 ID AAG43662
 XX AAG43662 standard; Protein; 285 AA.
 AC AAG43662;

XX 18-OCT-2000 (first entry)
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 54599.
 XX Arabidopsis thaliana.
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX Arabidopsis thaliana.
 OS Arabidopsis thaliana.
 PN EPI033405-A2.
 XX 06-SEP-2000.
 XX 25-FEB-2000; 2000EP-0301439.
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 16-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 11-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.

PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 26-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161320.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 4.6%; Score 113; DB 21; Length 315;
 Best Local Similarity 23.4%; Pred. No. 0.0029;
 Matches 79; Conservative 35; Mismatches 127; Indels 96; Gaps 15;

QY 6 SGVDIEL--SLWEIYSPRSGWMDFTQHLGVCCLVALLISVGLLSVAACWFLPSIIIAAA 63
 Db 53 ngndvicddwvrcpcrthpgwilgfrvgcflla-----nniarfa 96

QY 64 S-----WILTCVLLCCKHARCFILLVFLSC--GLREGRNALIAAGTGVILGH 110
 Db 97 nrgwriyytqwtftlialyfgms-----llsiygclykqkgntgladqvgi----d 148
 QY 111 VENIFHNEKGLDGMTCNLRAKSFSTHFPLLRKKYIEAIOIYGLATPLSVFDDOLVSNQOT 170
 Db 149 aengfrs--plidgdnmvsfekrktsgsalksyhlfqiiygmgaagaavltidsiyw--t 204
 QY 171 LAVSLFSPSHVLEAQLNDSKGEVLSVLYQMATTEVLSLQKLLAFAGLSLVLLGTGLF 230
 Db 205 vifpfls-----lqd-----yemsfmtvnlht-----snlvlllidtf 237
 QY 231 MKRFLGPCGWKYENIYITRQFVQDERERHQORPCVPLNKERRRYVIIPFWPWPKER 290
 Db 238 lnrkfpifrsyfilwtgcfvlfq-----wil-----hmfisvqwpyp--f 277
 QY 291 KNLGLFFLPI-----LIHLCTIWLFAAVDYLLYRLI 321
 Db 278 lnlsldmapvwyllvallhlpsygfalivkikyli 314

RESULT 14
 AAY54070
 ID AAY54070 standard; Protein; 681 AA.
 XX
 AC AAY54070;
 XX
 DT 27-MAR-2000 (first entry)
 XX
 DE Enzyme EPS6 which is involved in exopolysaccharide biosynthesis.
 XX
 KW Exopolysaccharide; EPS; ESP enzyme; EPS1; EPS2; EPS3; EPS4; EPS5;
 KW EPS6; EPS7; EPS8; EPS9; EPS10; Streptococcus thermophilus strain Sfi139;
 KW activated D-galactose pyranose; saccharide; beta-glycosyltransferase;
 KW transporter; food; fermented milk product; yoghurt; cheese;
 KW flavour stability; organoleptic property.
 XX
 OS Streptococcus thermophilus.
 XX
 PN WO9962316-A2.
 XX
 PD 09-DEC-1999.
 XX
 PF 22-APR-1999; 99WO-EP02841.
 XX
 PR 22-APR-1998; 98EP-0201310.
 PR 22-APR-1998; 98EP-0201311.
 PR 22-APR-1998; 98EP-0201312.
 XX
 PA (NEST) SOC PROD NESTLE SA.
 XX
 PI Stingele F, Germond JE, Lamothe G;
 XX
 DR WPI; 2000-097267/08.
 DR N-PSDB; AAZ45258, AAY54072, AAY54073, AAY54074.
 XX
 PT New recombinant enzymes for synthesis of exopolysaccharides,
 PT particularly in lactic acid bacteria, for improving properties of
 PT fermented milk products -
 XX
 PS Claim 3; Page 102-104; 162pp; French.
 XX
 CC AAY54065-74 represent enzymes involved in the biosynthesis of
 CC exopolysaccharides (EPS). These enzymes are designated EPS1-EPS10. and
 CC are encoded by open reading frames eps1-eps10. The enzymes are isolated
 CC from Streptococcus thermophilus strain Sfi139. The proteins are used
 CC in a method for the synthesis of EPS, which includes at least one step
 CC of forming a bond (alpha or beta-isomer) between C-1 (carrying the
 CC reducing aldehyde function, of an activated D-galactose pyranose), and
 CC a phosphate on a lipophilic or proteinaceous primer. Stepwise synthesis
 CC of EPS occurs with, in each step, addition of a new sugar unit, through
 CC its hemi-acetyl function, to an alcoholic hydroxyl of a second sugar
 CC unit, present at the end of a chain of sugar residues bonded to the

CC primer. EPS1 to EPS4 are used to elongate the EPS chain, and to regulate
 CC EPS production. EPS5 creates new bonds between saccharides, EPS6 is used
 CC in the biosynthesis of EPS, EPS7 and EPS8 are beta-glycosyltransferases,
 CC EPS9 is a transporter of repetitive units, and EPS10 catalyses the
 CC conversion of a pyranose form of a beta-D-galactose to the furanose
 CC form. The EPS enzyme are used to improve properties of foods,
 CC particularly fermented milk products such as yoghurt and cheese,
 CC e.g. their organoleptic properties and flavour stability.
 XX Sequence 681 AA;
 SQ

Query Match 4.1%; Score 101; DB 21; Length 681;
 Best Local Similarity 21.0%; Pred. No. 0.16;
 Matches 82; Conservative 62; Mismatches 164; Indels 82; Gaps 15;

QY 117 NFKGLDGMTCNLRKSFSPHPLKK-YIEAIQWIY-----LATPLSV----- 160
 Db 93 nfdnildrtvleehlekfdilpkirhyietieshykthfekdliateeiikllypdy 152
 QY 161 FDDLVSNOTLAVSLFSPSHVLEAQNDSKGEVLSVLYQMATTTEVLSLGGKLLAFAGL 220
 Db 153 ldfyysalkrksahmfimdkyfnnycewlsilfelekvlidiseyspfarvgrv 212
 QY 221 SLVLLGTGLFMKRFGLPCGWKYNIIYITRQVFQDERERHQORPCVLPNKEERKYYVII 280
 Db 213 seilldvwifk-----nnlnfteipvmfmeqgnw-----dkskrfisa 251
 QY 281 PTFWPTPKERNKLGFLFPLILHLICWVLFRAVDYLLYLIFSVSKQFOSLPGFEV---- 336
 Db 252 klnfkyytrrnklltigilli-----ifmtifdyihktvfpvmfmslflilisls 306
 QY 337 ----HLKLGHEKQGTQDIHDSFNISVPEPNC--IPKPKFL-----LSETWV 378
 Db 307 mrllynreysiksievivgmiffslgvf---ctrivshfknqnnvinyddnlnvnt 363
 QY 379 PLSVILLILVMLGLLSILMQLKILVSASFYPSVERKRIQYHLHAKLLKRSKQPLGEVKR 438
 Db 364 flk-illivttgnvfslfsgkflilggysylelrmilgyngaepl---itnplvni-- 417
 QY 439 RLSLYLT-----KIHFWLPVKLMIRKKOM 462
 Db 418 -ltryisgpgltalipf--sifflirrkni 444

RESULT 15
 AAY43772
 ID AAY43772 standard; Protein; 681 AA.
 XX
 AC AAY43772;
 XX
 DT 11-FEB-2000 (first entry)
 XX
 DE Amino acid sequence of eps6 of Streptococcus thermophilus Sfi39.
 XX
 KW eps operon; Streptococcus thermophilus Sfi39; enzyme; eps1; eps2;
 KW eps3; eps4; eps5; eps7; eps8; eps9; eps10;
 KW exopolysaccharide biosynthesis; EPS; intersugar bond; antitumour;
 KW probiotic; foodstuff; organoleptic quality; flavour;
 KW lactic acid bacteria; acidified milk product; yoghurt; cheese.
 XX
 OS Streptococcus thermophilus.
 XX
 PN W09954475-A2.
 XX
 PD 28-OCT-1999.
 XX
 PF 22-APR-1999; 99WO-EP03011.
 XX
 PR 22-APR-1998; 98EP-0201310.
 PR 22-APR-1998; 98EP-0201311.
 PR 22-APR-1998; 98EP-0201312.
 XX

PA (NEST) SOC PROD NESTLE SA.
 XX
 PI Stingele F, Germond JE, Lamothe G;
 XX
 DR WPI; 2000-013255/01.
 DR N-PSDB; AAZ30355, AAY43774, AAY43775, AAY43776;
 XX
 PT New recombinant enzymes for biosynthesis of exopolysaccharides having
 PT e.g. antitumor or probiotic properties or useful in fermented milk
 PT products
 XX
 PS Claim 3; Page 103-105; 163pp; French.
 XX
 CC AAY43767-76 represent the proteins encoded by the eps operon of
 CC Streptococcus thermophilus Sfi39. The operon contains 10 open reading
 CC frames, and encodes enzymes (eps1, eps2, eps3, eps4, eps5, eps6, eps7,
 CC eps8, eps9 and eps10) that are involved in the biosynthesis of
 CC exopolysaccharides (EPS). The enzymes catalyse the formation of
 CC specific intersugar bonds. The enzymes catalyse a process that includes
 CC at least one step of forming a bond (in alpha or beta anomeric form)
 CC between Cl, carrying the reducing aldehyde group of an activated D-Galp
 CC (galactose in pyranose form), and a phosphate group on a lipophilic or
 CC proteinaceous primer. The enzymes are used to produce EPS that have
 CC antitumor or probiotic properties or are used in foodstuffs to improve
 CC organoleptic qualities and flavour. When expressed by lactic acid
 CC bacteria, EPS impart a free-flowing character and/or a smooth, creamy
 CC texture to acidified milk products (yoghurt or cheese).
 XX
 SQ Sequence 681 AA;
 Query Match 4.1%; Score 101; DB 21; Length 681;
 Best Local Similarity 21.0%; Pred. No. 0.16;
 Matches 82; Conservative 62; Mismatches 164; Indels 82; Gaps 15;

QY 117 NFKGLDGMTCNLRKSFSPHPLKK-YIEAIQWIY-----LATPLSV----- 160
 Db 93 nfdnildrtvleehlekfdilpkirhyietieshykthfekdliateeiikllypdy 152
 QY 161 FDDLVSNOTLAVSLFSPSHVLEAQNDSKGEVLSVLYQMATTTEVLSLGGKLLAFAGL 220
 Db 153 ldfyysalkrksahmfimdkyfnnycewlsilfelekvlidiseyspfarvgrv 212
 QY 221 SLVLLGTGLFMKRFGLPCGWKYNIIYITRQVFQDERERHQORPCVLPNKEERKYYVII 280
 Db 213 seilldvwifk-----nnlnfteipvmfmeqgnw-----dkskrfisa 251
 QY 281 PTFWPTPKERNKLGFLFPLILHLICWVLFRAVDYLLYLIFSVSKQFOSLPGFEV---- 336
 Db 252 klnfkyytrrnklltigilli-----ifmtifdyihktvfpvmfmslflilisls 306
 QY 337 ----HLKLGHEKQGTQDIHDSFNISVPEPNC--IPKPKFL-----LSETWV 378
 Db 307 mrllynreysiksievivgmiffslgvf---ctrivshfknqnnvinyddnlnvnt 363
 QY 379 PLSVILLILVMLGLLSILMQLKILVSASFYPSVERKRIQYHLHAKLLKRSKQPLGEVKR 438
 Db 364 flk-illivttgnvfslfsgkflilggysylelrmilgyngaepl---itnplvni-- 417
 QY 439 RLSLYLT-----KIHFWLPVKLMIRKKOM 462
 Db 418 -ltryisgpgltalipf--sifflirrkni 444

Search completed: January 31, 2002, 08:01:45
 Job time: 66 sec